

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:50:39 ; Search time 73.8459 Seconds
(without alignments)
518.502 Million cell updates/sec

Title: US-10-674-755-11

Perfect score: 485

Sequence: 1 LKIDSDSDYVKGFRAP.....KXAELEKTRDLKKAHVPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	485	100.0	204	2 AAW14571	Aaw14571 Streptococ
2	485	100.0	204	7 ABW02605	Abw02605 Ef1019c p
3	485	100.0	8991	6 ABU08487	Abu08487 S. pneumo
4	469	96.7	198	7 ABW02615	Abw02615 RxlC pneu
5	469	96.7	315	2 AAY04375	Aay04375 Streptoco
6	469	96.7	619	2 AAR63437	Aar63437 Streptoco
7	469	96.7	619	2 AAR87598	Aar87598 Pneumococ
8	469	96.7	619	2 AAR86911	Aar86911 Pneumococ
9	469	96.7	619	2 AAY41838	Aay41838 Streptoco
10	469	96.7	619	5 AAE18782	Aae18782 S. pneumo
11	469	96.7	619	6 ABU45778	Abu45778 Protein e
12	469	96.7	619	8 ADOS2126	Ado52126 Streptoco
13	469	96.7	648	2 AAW70336	Aaw70336 Pneumococ
14	469	96.7	648	2 AAW62274	Aaw62274 Streptoco
15	469	96.7	648	2 AAY41837	Aay41837 Streptoco
16	469	96.7	648	2 AAW87879	Aaw87879 A pneumoc
17	469	96.7	653	2 AAW92456	Aaw92456 S. pneumo
18	469	96.7	684	2 AAR73912	Aar73912 Streptoco
19	469	96.1	198	2 AAW14581	Aaw14581 Streptoco
20	449	92.6	653	2 AAR27150	Aar27150 PspA frag
21	446.5	92.1	289	2 AAW62276	Aaw62276 Streptoco
22	446.5	92.1	289	2 AAY41840	Aay41840 Streptoco
23	446.5	92.1	289	2 AAW87910	Aaw87910 Protein s
24	446.5	92.1	289	2 AAW92458	Aaw92458 S. pneumo
25	442	91.1	195	2 AAW14591	Aaw14591 Streptoco

26	442	91.1	195	7 ABW02625	Abw02625 Wu2c pneu
27	423	87.2	623	6 ABU08494	Abu08494 Fragment
28	414	85.4	170	7 ABW02614	Abw02614 Rct135c p
29	414	85.4	181	7 ABW02596	Abw02596 0921134c
30	414	85.4	865	6 ABU08489	Abu08489 S. pneumo
31	414	85.4	929	2 AAW14593	Aaw14593 Streptoco
32	414	85.4	929	2 AAY43384	Aay43384 S. pneumo
33	411	84.7	188	2 AAW14580	Aaw14580 Streptoco
34	411	84.7	188	7 ABW02613	Abw02613 Rct129c p
35	410	84.5	1231	6 ABU08490	Abu08490 Fragment
36	402	82.9	588	6 ABU08491	Abu08491 Coiled co
37	402	82.9	589	2 AAY43392	Aay43392 PspC alph
38	400	82.5	204	2 AAW14578	Aaw14578 Streptoco
39	400	82.5	204	7 ABW02612	Abw02612 Rct123c p
40	399.5	82.4	180	2 AAW14562	Aaw14562 Streptoco
41	396.5	81.8	187	2 AAW14579	Aaw14579 Streptoco
42	389	80.2	206	2 AAW14574	Aaw14574 Streptoco
43	389	80.2	206	7 ABW02608	Abw02608 DB15C pne
44	357.5	73.7	550	8 ADK48356	Adk48356 Streptoco
45	357.5	73.7	550	8 ADR95223	Adr95223 Novel S.

ALIGNMENTS

RESULT 1
AAW14571
ID AAW14571 standard; protein; 204 AA.

AC AAW14571;
XX
XX
DT 17-OCT-2003 (revised)
DT 28-OCT-1997 (first entry)
XX
DE Streptococcus pneumoniae PspA central region.

XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW bacteraemia; pneumonia.
XX
OS Streptococcus pneumoniae; strain Ef1019.

XX
XX
PN WO9709994-Al.
XX
PD 20-MAR-1997.

XX
PF 16-SEP-1996; 96WO-US014819.
XX
PR 15-SEP-1995; 95US-00529055.

XX (UABR-) UAB RES FOUND.
XX Briles DB, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX WPI; 1997-202002/18.

XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
in vaccines for protecting animals against S.pneumoniae infection.

PS Example 6; Fig 13; 296pp; English.

XX This sequence shows the central portion, including the C-terminus of the
alpha-helix region and some of the proline-rich region, of pneumococcal
surface protein A (PspA) of Streptococcus pneumoniae strain Ef1019.

CC Comparison of the N-terminal and central regions (AAW14533-57 and
AAW14562-91) of PspA polypeptides from different pneumococcal strains can
be used to divide the strains into several families based on sequence
homologies. PspA polypeptides, or fragments of them, can be used in
vaccines to protect animals against S. pneumoniae infection and hence for
the prevention of diseases such as otitis media, meningitis, bacteraemia
and pneumonia. The sequence of the 3' half of the PspA alpha-helical
region and the immediate 5' tip of the coding sequence are likely to be
the critical sequences for predicting PspA cross-reactions and vaccine

```

CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 204 AA;

Query Match      100.0%; Score 485; DB 2; Length 204;
Best Local Similarity 100.0%; Pred. No. 9.2e-37;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLELDQKAAE 60
Db 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLELDQKAAE 60
Qy 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKKA VNEPE 99
Db 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKKA VNEPE 99

RESULT 2
ABW02605
ID ABW02605 standard; protein; 204 AA.
XX
AC ABW02605;
XX
XX 12-FEB-2004 (first entry)
XX
DE Ef1019c pneumococcal surface protein A (PspA) central region.
XX
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
KW immunological; gene therapy; immunostimulant.
XX
OS Unidentified.
XX
XX US6592876-B1.
XX
PD 15-JUL-2003.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX 20-APR-1993; 93US-00048896.
XX 06-JUN-1995; 95US-00465746.
XX
XX (UABR-) UAB RES FOUND.
XX
PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX
XX WPI; 2003-862841/80.
XX
XX Immunological composition for obtaining expression products used for
PT detecting the presence of Streptococcus pneumoniae or its strain,
PT comprises at least two different full length isolated gene encoding
PT pneumococcal surface protein A.
XX
XX Example 6; SEQ ID NO 51; 121pp; English.
XX
XX The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspAs) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antigenic,
CC immunological or vaccine compositions, for eliciting antibodies, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Ef1019c
CC pneumococcal surface protein A (PspA) central region. This sequence is
CC used in the exemplification of the invention
XX
XX Sequence 204 AA;

Query Match      100.0%; Score 485; DB 7; Length 204;
Best Local Similarity 100.0%; Pred. No. 9.2e-37;

Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLELDQKAAE 60
Db 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLELDQKAAE 60
Qy 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKKA VNEPE 99
Db 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKKA VNEPE 99

us-10-674-755-11.rag
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLELDQKAAE 60
Db 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLELDQKAAE 60
Qy 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKKA VNEPE 99
Db 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKKA VNEPE 99

RESULT 3
ABU08487
ID ABU08487 standard; protein; 8991 AA.
XX
AC ABU08487;
XX
XX 24-JUN-2003 (first entry)
XX
DE S. pneumoniae pneumococcal surface protein A (PspA) protein.
XX
KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
KW antibacterial.
XX
OS Streptococcus pneumoniae.
XX
XX Key Location/Qualifiers
FH Misc-difference 1. .8991
FT /note= "All Xaa residues within this sequence are
FT unknown"
XX
XX US6500613-B1.
XX
XX 31-DEC-2002.
XX
XX 16-SEP-1996; 96US-00714741.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX (UYAL-) UNIV ALABAMA.
XX
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
XX Hollingshead S, Tart R, Brooks-Walter A;
XX
XX WPI; 2003-361534/34.
XX
XX Isolated PspC amino acid sequence used as polymerase chain reaction or
PT hybridization probe, comprises pneumococcal surface protein having alpha-
PT helical, proline rich and repeat regions.
XX
XX Disclosure; Col 145-188; 186pp; English.
XX
XX The present invention relates to the isolation of Streptococcus
CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-
CC like protein having alpha-helical, proline rich and repeat regions. The
CC PspC and PspA proteins may be used in a vaccine to protect against
CC pneumococcal infections. The polynucleotide sequences encoding PspC and
CC PspA may be used for the expression of the proteins, and as PCR primers
CC or hybridisation probes. The present sequence represents S. pneumoniae
CC PspA protein
XX
XX Sequence 8991 AA;

Query Match      100.0%; Score 485; DB 6; Length 8991;
Best Local Similarity 100.0%; Pred. No. 1.1e-34;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLELDQKAAE 60
Db 5322 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLELDQKAAE 5381

```

QY 61 ENNVEDYFKGLEKTIAAKAELEKTEADLKAVNEPE 99
 DB 5382 ENNVEDYFKGLEKTIAAKAELEKTEADLKAVNEPE 5420

RESULT 4

ABW02615
 ID ABW02615 standard; protein; 198 AA.

XX AC ABW02615;
 XX DT 12-FEB-2004 (first entry)
 XX DE RxlC pneumococcal surface protein A (PspA) central region.
 XX KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
 XX KW immunological; gene therapy; immunostimulant.
 XX OS Unidentified.
 XX PN US6592876-B1.
 XX PD 15-JUL-2003.
 XX PF 15-SEP-1995; 95US-00529055.
 XX PR 20-APR-1993; 93US-00048896.
 XX PR 06-JUN-1995; 95US-00465746.
 XX PA (UABR-) UAB RES FOUND.

XX PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooke-Walter A;
 XX DR WPI; 2003-862841/80.

XX PT Immunological composition for obtaining expression products used for
 PT detecting the presence of Streptococcus pneumoniae or its strain,
 PT comprises at least two different full length isolated gene encoding
 PT pneumococcal surface protein A.

XX PS Example 6; SEQ ID NO 61; 121pp; English.

XX CC The present invention relates to an immunological composition comprising
 CC at least 2 different full length isolated genes encoding pneumococcal
 CC surface protein A (PspAs) from different groups based on restriction
 CC fragment polymorphism analysis. The invention is useful for obtaining
 CC expression products by recombinant techniques to detect, determine,
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its
 CC strain. The expression product is useful for preparing antigenic,
 CC immunological or vaccine compositions, for eliciting antibodies, an
 CC immunological response (other than or additional to antibodies) or a
 CC protective response (including antibody or other immunological response
 CC by administering compositions to a host). The invention is also useful as
 CC vaccines and in gene therapy. The present sequence is RxlC pneumococcal
 CC surface protein A (PspA) central region. This sequence is used in the
 CC exemplification of the invention

XX SQ Sequence 198 AA;

Query Match 96.7%; Score 469; DB 7; Length 198;
 Best Local Similarity 96.0%; Pred. No. 2.6e-35;
 Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKIDSDSDYVKGFPAPLQSELDKAKLSELSKIDELDAETAKLEDQKAAE 60
 DB 1 LKIDSESESDYAKGFPAPLQSKLDKAKLSELSKIDELDAETAKLEDQKAAE 60

QY 61 ENNVEDYFKGLEKTIAAKAELEKTEADLKAVNEPE 99
 DB 61 ENNVEDYFKGLEKTIAAKAELEKTEADLKAVNEPE 99

RESULT 5

AAV04375
 ID AAV04375 standard; protein; 315 AA.
 XX AC AAV04375;

XX DT 23-JUN-1999 (first entry)
 XX DE Streptococcus pneumoniae PspA protein sequence.

XX KW Streptococcus pneumoniae; pspA; pneumococcal; surface protein; vaccine;
 XX KW immunological; infection.
 XX OS Streptococcus pneumoniae.
 XX OS Synthetic.

XX PN WO9914333-A2.

XX PD 25-MAR-1999.

XX PF 18-SEP-1998; 98WO-US019740.

XX PR 18-SEP-1997; 97US-00932982.

XX PA (INMR) PASTEUR MERIEUX CONNAUGHT.

XX PI Becker R, Gray M, Pyle D;

XX WPI; 1999-229537/19.

XX DR N-PSDB; AAX33124.

XX PT DNA encoding PspA molecule with modified internal translational
 PT initiation sites.

XX PS Disclosure; Page; 36pp; English.

XX CC The present sequence represents a pneumococcal surface protein A (PspA)
 CC molecule where internal naturally occurring translational initiation
 CC sites have been modified or eliminated so that expression of the DNA
 CC sequence results in a single form of PspA. The PspA nucleotide sequence
 CC can be used to transform a unicellular host to produce the PspA protein.
 CC The PspA protein can be used in an immunological composition for treating
 CC or preventing S. pneumoniae infection especially in a child. Antibodies
 CC to the PspA protein can also be used to treat S. pneumoniae infection.
 CC The immunogenic peptides are designed to confer broad protection against
 CC diverse pneumococcal strains. N.B. The present sequence is not given in
 CC the specification but is encoded by the sequence given in AAX33124

XX SQ Sequence 315 AA;

Query Match 96.7%; Score 469; DB 2; Length 315;
 Best Local Similarity 96.0%; Pred. No. 4.7e-35;
 Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKIDSDSDYVKGFPAPLQSELDKAKLSELSKIDELDAETAKLEDQKAAE 60
 DB 193 LKIDSESESDYAKGFPAPLQSKLDKAKLSELSKIDELDAETAKLEDQKAAE 252

QY 61 ENNVEDYFKGLEKTIAAKAELEKTEADLKAVNEPE 99

DB 253 ENNVEDYFKGLEKTIAAKAELEKTEADLKAVNEPE 291

RESULT 6

AAV63437
 ID AAV63437 standard; protein; 619 AA.

XX AC AAV63437;

XX DT 09-SEP-2004 (revised)
 DT 16-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 19-JUL-1995 (first entry)

XX

```

DE Pneumococcal surface protein A from S.pneumoniae Rxl.
XX
XX Pneumococcal surface protein A; PspA; Streptococcus; PCR; pneumococcal;
KW primer; protection-eliciting epitope; epitope; vaccine; amplify.
XX
XX Streptococcus pneumoniae.
OS Unidentified.
XX
XX Key Location/Qualifiers
XX Protein 192..260
FT /label= "protein fragment of Claim 1"
FT
XX
XX EP622081-A2.
XX
XX 02-NOV-1994.
XX
XX 19-APR-1994; 94EP-00302767.
XX
XX 20-APR-1993; 93US-00048896.
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, Yother JL, McDaniel LS;
XX WPI; 1994-359522/45.
XX N-PSDB; AAQ78131.
XX
XX regions of Pneumococcal surface protein A - derived from the Rxl PspA
XX strain, for the preparation of cross-reactive vaccines for the prevention
XX of pneumococcal infections.
XX
XX Disclosure; Page 13-16; 26pp; English.
XX
XX The amino acid sequence of the novel Pneumococcal surface protein A
XX (PspA) from Streptococcus pneumoniae strain Rxl. The gene was PCR
XX amplified from S.pneumoniae genomic DNA using the primers AAQ78132-5. The
XX gene was used to derive truncated peptide fragments containing protection
XX -eliciting epitopes for use in vaccines against pneumococcal diseases.
XX The epitopic fragments are derived from amino acids 192-260 and
XX optionally contain a further 25 a.a. residues at both the N- and C-
XX terminal regions of the peptide. The epitopic peptide fragments may be
XX derived from different strains of S.pneumoniae and are homologous to the
XX Rxl strain epitope. (Updated on 25-MAR-2003 to correct PN field.)
XX (Updated on 16-OCT-2003 to standardise OS field)
XX
XX Revised record issued on 09-SEP-2004 : Correction to feature table key
XX
XX Sequence 619 AA;
XX
Query Match 96.7%; Score 469; DB 2; Length 619;
Best Local Similarity 96.0%; Pred. No. 1.1e-34;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 1 LKEIDESDSEYVKEGFRAPLQSLDAKQAKLSKLELSDKIDELDAEIAKLSDQLKAAE 60
Db 223 LKEIDESDSEYVKEGFRAPLQSLDAKQAKLSKLELSDKIDELDAEIAKLSDQLKAAE 282
Qy 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKKAVNEPE 99
Db 283 ENNVEDYFKEGLEKTTAAKKAELKTEADLKKAVNEPE 321
RESULT 7
AAR87598
ID AAR87598 standard; protein; 619 AA.
AC AAR87598;
XX
XX 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 04-JUL-1996 (first entry)
XX
XX Pneumococcal surface protein (PspA) .
DE

```

```

XX PspA; pneumococcal surface protein; truncated; immunoprotective;
KW soluble fragment; insertion-duplication mutagenesis.
XX
XX Streptococcus pneumoniae; strain Rxl.
OS
XX
XX Key Location/Qualifiers
XX Peptide 1..31
FT /label= signal_peptide
FT Protein 32..619
FT /label= mature_protein
FT Region 32..319
FT /label= alpha-helical coiled-coil region
FT /note= "contains a seven-residue periodicity"
FT
FT Region 320..401
FT /note= "proline-rich region"
FT Region 402..421
FT /note= "repeat region"
FT Region 422..441
FT /note= "repeat region"
FT Region 442..461
FT /note= "repeat region"
FT Region 462..481
FT /note= "repeat region"
FT Region 482..501
FT /note= "repeat region"
FT Region 502..521
FT /note= "repeat region"
FT Region 522..541
FT /note= "repeat region"
FT Region 542..561
FT /note= "repeat region"
FT Region 562..581
FT /note= "repeat region"
FT Region 582..619
FT /note= "hydrophobic region starts in last repeat region
is potential membrane-spanning region"
XX
XX US5476929-A.
XX
XX 19-DEC-1995.
XX
XX 03-JUN-1993; 93US-00072070.
XX
XX 15-FEB-1991; 91US-00656773.
XX 12-FEB-1992; 92US-00835698.
XX
XX (UABR-) UAB RES FOUND.
XX
XX McDaniel LS, Yother JL, Briles DE;
XX WPI; 1996-049021/05.
XX N-PSDB; AAT08979.
XX
XX New pneumococcal surface protein A fragments - comprise proline-rich
XX region and/or repeat region, used for detection of Streptococcus
XX pneumoniae.
XX
XX Claim 1; Col 15-20; 23pp; English.
XX
XX The present sequence is that of PspA (pneumococcal surface protein A)
XX encoded by AAT08979. Through the technique of insertion-duplication
XX mutagenesis of the PspA gene of the strain Rxl of Streptococcus
XX pneumoniae with plasmids contg. cloned fragments of the PspA structural
XX gene, it has been possible to produce soluble fragments of PspA that are
XX secreted by pneumococci. The method can be used to provide an
XX immunoprotective truncated PspA protein. Primers and probes based on the
XX present sequence are claimed, and are useful for the detection of (at
XX least 32) S. pneumoniae strains. (Updated on 25-MAR-2003 to correct PF
XX field.) (Updated on 16-OCT-2003 to standardise OS field)
XX
XX Sequence 619 AA;
XX

```


Matches	95;	Conservative	3;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	1	LKEIDESDS	YVKEGFRAP	QSQELDAKQAKLSKLEELSDKIDELDAEIAKLESDQLKAAE	60				
Db	223	LKEIDSESD	YAKEGFRAP	QSQELDAKQAKLSKLEELSDKIDELDAEIAKLESDQLKAAE	282				
Qy	61	ENNNVDY	FKEGLEKTI	AAKKAELKTEADLKKAVNEPE	99				
Db	283	ENNNVDY	FKEGLEKTI	AAKKAELKTEADLKKAVNEPE	321				
RESULT 9									
AAAY41838									
ID	AAAY41838	standard; protein; 619	AA.						
XX	XX	AAAY41838;							
XX	XX								
DT	08-DEC-1999	(first entry)							
XX	XX								
DE									
XX	XX	Streptococcus pneumoniae	Rx1 PspA protein sequence.						
XX	XX								
KW	KW	Streptococcus pneumoniae	Rx1; PspA; immunoprotective; vaccine; diagnosis;						
KW	KW	infection; pneumococcal surface protein A.							
XX	XX								
OS	OS	Streptococcus pneumoniae.							
XX	XX								
PN	PN	US5965400-A.							
XX	XX								
PD	PD	12-OCT-1999.							
XX	XX								
PF	PF	23-MAY-1994; 94US-00247491.							
XX	XX								
PR	PR	15-FEB-1991; 91US-00656773.							
XX	XX								
PR	PR	12-FEB-1992; 92US-00835698.							
XX	XX								
PA	PA	(UABR-) UAB RES FOUND.							
XX	XX								
PI	PI	Yother JL, Briles DE;							
XX	XX								
DR	DR	WPI; 1999-579913/49.							
DR	DR	N-PSDB; AAZ25063.							
XX	XX								
PT	PT	DNA encoding a truncated pneumococcal surface protein A used in the							
PT	PT	development of pneumococcal infections.							
XX	XX								
PS	PS	Claim 1; Fig 3; 27pp; English.							
XX	XX								
CC	CC	The present sequence represents Streptococcus pneumoniae Rx1							
CC	CC	immunoprotective Pneumococcal surface protein A (PspA). The present							
CC	CC	invention also describes a method of forming the immunoprotective							
CC	CC	truncated PspA, comprising incorporating a vector comprising the isolated							
CC	CC	DNA molecule encoding PspA (I), into a bacterium via transformation. (I)							
CC	CC	is used to design primers which are capable of detecting a large number							
CC	CC	of S. pneumoniae strains, which in turn can be used to diagnose							
CC	CC	pneumococcal infection in mammals (e.g. humans), independent of the							
CC	CC	strain which has caused it. The PspA protein is used to develop a vaccine							
CC	CC	against pneumococcal infection comprising, as an immunologically-active							
CC	CC	component, a live attenuated or killed bacteria containing a gene coding							
CC	CC	for the truncated form of PspA							
XX	XX								
SQ	SQ	Sequence 619	AA;						
		Query Match	96.7%;	Score	469;	DB	2;	Length	619;
		Best Local Similarity	96.0%;	Pred. No.	1.1e-34;				
		Matches	95;	Conservative	3;	Mismatches	1;	Indels	0;
		Gaps	0;						
Qy	1	LKEIDESDS	YVKEGFRAP	QSQELDAKQAKLSKLEELSDKIDELDAEIAKLESDQLKAAE	60				

```

RESULT 10
AAE18782
ID AAE18782 standard; protein; 619 AA.
XX
AC AAE18782;
XX
XX 17-MAY-2002 (first entry)
XX
XX S. pneumoniae Rx1 strain pneumococcal surface protein A (PspA).
XX
XX Coiled-coil structural scaffold; heptad repeat; epitope; immune response;
XX cell-mediated immunity; microbial infection; cross-protection; therapy;
XX antimicrobial; vaccine; pneumococcal surface protein A; PspA.
XX
XX Streptococcus pneumoniae.
XX
FH Key Location/Qualifiers
FT Domain 1..314
FT /label= Helical_domain
FT Region 1..303
FT /note= "N-terminal region"
FT Region 38..44
FT /note= "Immunogenic region 3"
FT Region 40..46
FT /note= "Immunogenic region 5"
FT Region 75..80
FT /note= "Immunogenic region 29"
FT Region 82..87
FT /note= "Immunogenic region 52"
FT Region 96..101
FT /note= "Immunogenic region 66"
FT Region 114..119
FT /note= "Immunogenic region 73"
FT Region 130..135
FT /note= "Immunogenic region 78"
FT Region 137..142
FT /note= "Immunogenic region 89"
FT Region 140..145
FT /note= "Immunogenic region 91"
FT Region 152..156
FT /note= "Immunogenic region 95"
FT Domain 153..170
FT /label= Coiled_coil_motif
FT Region 161..164
FT /note= "Immunogenic region 101"
FT Region 166..170
FT /note= "Immunogenic region 116"
FT Region 173..177
FT /note= "Immunogenic region 122"
FT Region 176..180
FT /note= "Immunogenic region 123"
FT Domain 181..198
FT /label= Coiled_coil_motif
FT Region 187..191
FT /note= "Immunogenic region 130"
FT Region 194..198
FT /note= "Immunogenic region 133"
FT Region 215..219
FT /note= "Immunogenic region 140"
FT Region 226..230
FT /note= "Immunogenic region 145"
FT Region 229..232
FT /note= "Immunogenic region 148"
FT Region 260..264
FT /note= "Immunogenic region 166"
FT Region 263..267
FT /note= "Immunogenic region 168"
FT Region 284..287
FT /note= "Immunogenic region 76"
FT Region 286..290
FT /note= "Immunogenic region 179"
FT

Region 294..298
/Note= "Immunogenic region 182"
Region 295..299
/Note= "Immunogenic region 185"
Region 309..313
/Note= "Immunogenic region 195"
Region 322..326
/Note= "Immunogenic region 206"
XX
XX WO200196368-A2.
XX
XX 20-DEC-2001.
XX
XX 14-JUN-2001; 2001WO-US019168.
XX
XX 14-JUN-2000; 2000US-0211892P.
XX
XX 23-JUN-2000; 2000US-0213387P.
XX
XX (CYTO-) CYTOVAX BIOTECHNOLOGIES INC.
XX
XX Houston ME, Hodges RS;
XX
XX WPI; 2002-188298/24.
XX
XX New synthetic peptide derived from naturally occurring microbial and non-
XX microbial protein antigen useful to stimulate and elicit an immune
XX response in an animal.
XX
XX Example 1; Page 90-92; 99pp; English.
XX
XX The invention relates to the uses of coiled-coil structural scaffold to
XX generate structure-specific peptides, including synthetic peptides
XX derived from naturally occurring microbial and non-microbial protein
XX antigens. The structure of the synthetic peptides utilizes a scaffold of
XX heptad repeat units into which epitopes derived from coiled-coil regions
XX of native proteins are spliced. The resulting peptide has a more stable
XX coiled-coil structure, hence improving presentation of the epitopes in a
XX helical conformation. The peptides of the invention are used to stimulate
XX and elicit an immune response in an animal, as vaccine, to treat or
XX prevent microbial infection by several strains and/or species of
XX microorganism, to provide cross-protection to at least one strain and/or
XX species of microorganism and to stimulate antibody production or cell-
XX mediated immunity to the naturally occurring protein. The present
XX sequence is Streptococcus pneumoniae Rx1 strain pneumococcal surface
XX protein A (PspA) which adopts a coiled-coil structure
XX
XX Sequence 619 AA;
XX
Query Match 96.7%; Score 469; DB 5; Length 619;
Best Local Similarity 96.0%; Pred. No. 1.1e-34;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 LKEIDSDSDYVKEGFRAPLOSELDAKQAKLSKLSBELSKIDELDAEIAKLDELQKAAE 60
DB 223 LKEIDSESDYAKEGFRAPLOSKLDKAKKLSKLSBELSKIDELDAEIAKLDELQKAAE 282
QY 61 ENNVEDYFKGLEKTIAAKKAELKTEADLKAVNEPE 99
DB 283 ENNVEDYFKGLEKTIAAKKAELKTEADLKAVNEPE 321
XX
XX RESULT 11
XX ABA45778
XX ID ABA45778 standard; protein; 619 AA.
XX
XX ABA45778;
XX
XX 19-JUN-2003 (first entry)
XX
XX Protein encoded by Prokaryotic essential gene #31305.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX

```

```

OS Streptococcus pneumoniae.
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US000107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
XX N-PSDB; ACA49648.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 73702; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed, (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than S. aureus, S. typhimurium,
XX K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
XX the target prokaryotic essential genes. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 619 AA;
XX
XX Query Match 96.7%; Score 469; DB 6; Length 619;
XX Best Local Similarity 96.0%; Pred. No. 1.1e-34;
XX Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 LKEIDSDSDYVKEGFRAPLQSELDKAKLQKLEELSDKIDELDAETAKLEDQKAAE 60
XX
XX 223 LKEIDSESDYAKGFRAPLQSKLDKAKKLEELSDKIDELDAETAKLEDQKAAE 282
XX
XX 61 ENNVEDYFKEGLEKTIKAAKAELEKTEADLKAVNEPE 99
XX
XX 283 ENNVEDYFKEGLEKTIKAAKAELEKTEADLKAVNEPE 321
XX
XX
XX RESULT 13
XX AAW70336
XX ID AAW70336 standard; protein; 648 AA.
XX

```

```

AC AAW70336;
XX 18-NOV-1998 (first entry)
XX Pneumococcal surface protein A (PspA).
DE
XX Pneumococcal surface protein A gene; PspA; PspA epitope; vaccine;
KW insertion-duplication mutagenesis; otitis media; meningitis; bacteraemia;
KW pneumonia.
XX Streptococcus pneumoniae.
XX OS
XX FH Key Location/Qualifiers
XX FT Peptide 1..31
XX FT Protein /note= "Signal peptide"
XX FT /note= "PspA"
XX FT Region 32..648
XX FT /note= "PspA"
XX FT Misc-difference 647
XX FT /note= "alpha-helical coil region representing the
XX truncated PspA of the invention"
XX FT /note= "Encoded by AGG"
XX PN US5804193-A.
XX PD 08-SEP-1998.
XX PF 17-MAR-1994; 94US-00214222.
XX PR 15-FEB-1991; 91US-00656773.
XX PR 12-FEB-1992; 92US-00835698.
XX PA (UABR-) UAB RES FOUND.
XX PI Briles DE, Yother JL;
XX WPI; 1998-505588/43.
XX DR N-PSDB; AAV33264.
XX Truncated pneumococcal surface protein - useful in vaccines against
XX pneumococcal infection.
XX Example 3; Fig 3A-3C; 22pp; English.
XX The present sequence represents the Streptococcus pneumoniae Rx1
XX pneumococcal surface protein A (PspA). The invention provides a purified
XX truncated form of PspA, formed by an insertion-duplication mutagenesis
XX technique, comprising of the first 288 N-terminal residues of the mature
XX form of wild-type PspA (AAW70336). The truncated PspA contains
XX immunoprotective epitopes of PspA. The invention claims for a vaccine
XX against pneumococcal infection, comprising live-attenuated or killed S.
XX pneumoniae, containing the gene coding for the truncated PspA protein.
XX The truncated protein, optionally conjugated to a poorly immunogenic or
XX nonimmunogenic molecule, is claimed to be useful in vaccines against
XX pneumococcal infection, especially otitis media, meningitis, bacteraemia
XX and pneumonia
XX
XX Sequence 648 AA;
Query Match 96.7%; Score 469; DB 2; Length 648;
Best Local Similarity 96.0%; Pred. No. 1.2e-34;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 LKEIDSDSDYKVEGFAPLQSELDKAKLQKLEELSDKIDELDAEIAKLEDLKAAE 60
DB 223 LKEIDSESDYKVEGFAPLQSKLDKAKLQKLEELSDKIDELDAEIAKLEDLKAAE 282
QY 61 ENNVEDYFKEGLEKTIKAAKAELEKTEADLKKA VNEPE 99
DB 283 ENNVEDYFKEGLEKTIKAAKAELEKTEADLKKA VNEPE 321
RESULT 14
AAY41837
ID AAY41837 standard; protein; 648 AA.

```

```

AAW62274
ID AAW62274 standard; protein; 648 AA.
XX
XX AC AAW62274;
XX DT 22-SEP-1998 (first entry)
XX DE Streptococcus pneumoniae pspA protein.
XX KW Streptococcus pneumoniae strain Rxi; pspA; immunoprotective; immunogen;
XX KW pneumococcal surface protein A; cholera toxin B subunit; fusion protein;
XX KW antigenic.
XX OS Streptococcus pneumoniae.
XX PN US5753463-A.
XX PD 19-MAY-1998.
XX PF 06-JUN-1995; 95US-00469434.
XX PR 15-FEB-1991; 91US-00656773.
XX PR 12-FEB-1992; 92US-00835698.
XX PR 03-JUN-1993; 93US-00072065.
XX PA (UABR-) UAB RES FOUND.
XX PI Yother JL, Briles DE;
XX WPI; 1998-311399/27.
XX DR N-PSDB; AAV39470.
XX Truncated pneumococcal surface protein and cholera toxin B sub-unit
XX fusion protein - useful as an immunogen against Streptococcus pneumoniae.
XX Claim 1; Fig 3; 22pp; English.
XX The present sequence represents the pneumococcal surface protein A (PspA)
XX protein from Streptococcus pneumoniae. A recombinant DNA molecule has
XX been developed which encodes a fusion protein comprising a truncated form
XX of PspA and cholera toxin B subunit (CTB), where the DNA molecule
XX comprises a nucleotide sequence encoding the truncated PspA linked by an
XX in-frame genetic fusion to a ctxB gene, and where the truncated PspA
XX contains immunoprotective epitopes and up to 90% of the whole PspA
XX protein, except for the cell membrane anchor region. The fusion protein
XX is useful for providing an immunogen to protect neonates and children
XX against S.pneumoniae. Most antigenic proteins of this strain are not
XX immunogenic enough to provide protection. The antigenic epitopes of the
XX fusion protein are directed against capsular polysaccharide antigens of
XX S.pneumoniae, specifically it contains the protective epitopes of PspA.
XX The protein can also be used in solid-phase immunoassay, since
XX it is readily bound to supports coated with monosialoganglioside GM1. The
XX fusion protein is more immunogenic against S.pneumoniae than using PspA
XX alone as the immunogen
XX
XX Sequence 648 AA;
Query Match 96.7%; Score 469; DB 2; Length 648;
Best Local Similarity 96.0%; Pred. No. 1.2e-34;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 LKEIDSDSDYKVEGFAPLQSELDKAKLQKLEELSDKIDELDAEIAKLEDLKAAE 60
DB 223 LKEIDSESDYKVEGFAPLQSKLDKAKLQKLEELSDKIDELDAEIAKLEDLKAAE 282
QY 61 ENNVEDYFKEGLEKTIKAAKAELEKTEADLKKA VNEPE 99
DB 283 ENNVEDYFKEGLEKTIKAAKAELEKTEADLKKA VNEPE 321
RESULT 15
AAY41837
ID AAY41837 standard; protein; 648 AA.

```

QY 61 ENNVEDYFKEGLEKTIAAKKALEKTEADLKKA VNEPE 99
DB 283 ENNVEDYFKEGLEKTIAAKKALEKTEADLKKA VNEPE 321

Search completed: June 21, 2005, 10:10:13
Job time : 75.8459 secs

XX AC AAY41837;
XX DT 08-DEC-1999 (first entry)
XX DE Streptococcus pneumoniae Rx1 PspA protein sequence.
XX KW Streptococcus pneumoniae Rx1; PspA; immunoprotective; vaccine; diagnosis;
XX KW infection; pneumococcal surface protein A.
XX OS Streptococcus pneumoniae.
XX PH Key Location/Qualifiers
FT Misc-difference 619. .620 /note= "a stop codon is present in the nucleotide
FT sequence at this position"
FT Misc-difference 621. .622 /note= "a stop codon is present in the nucleotide
FT sequence at this position"
FT Misc-difference 625. .626 /note= "a stop codon is present in the nucleotide
FT sequence at this position"
FT Misc-difference 630. .631 /note= "a stop codon is present in the nucleotide
FT sequence at this position"
FT Misc-difference 632. .633 /note= "a stop codon is present in the nucleotide
FT sequence at this position"
XX PN US5965400-A.
XX PD 12-OCT-1999.
XX PF 23-MAY-1994; 94US-00247491.
XX PR 15-FEB-1991; 91US-00656773.
XX PR 12-FEB-1992; 92US-00835698.
XX PA (UABR-) UAB RES FOUND.
XX PI Yother JL, Briles DE;
XX WPI; 1999-579913/49.
XX DR N-PSDB; AA225063.
XX PT DNA encoding a truncated pneumococcal surface protein A used in the
XX development of pneumococcal infections.
XX PS Claim 1; Fig 3; 27pp; English.
XX CC The present sequence represents Streptococcus pneumoniae Rx1
XX immunoprotective Pneumococcal surface protein A (PspA). The present
XX invention also describes a method of forming the immunoprotective
XX truncated PspA, comprising incorporating a vector comprising the isolated
XX DNA molecule encoding PspA (1), into a bacterium via transformation. (1)
XX is used to design primers which are capable of detecting a large number
XX of S. pneumoniae strains, which in turn can be used to diagnose
XX pneumococcal infection in mammals (e.g. humans), independent of the
XX strain which has caused it. The PspA protein is used to develop a vaccine
XX against pneumococcal infection comprising, as an immunologically-active
XX component, a live attenuated or killed bacteria containing a gene coding
XX for the truncated form of PspA
XX SQ Sequence 648 AA;

Query Match 96.7%; Score 469; DB 2; Length 648;
Best Local Similarity 96.0%; Pred. No. 1.2e-34;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSDKIDELDAETAKLEDQDKAAE 60
DB 223 LKEIDSESEDYAKEGFRAPLQSELDKQAKLSKLELSDKIDELDAETAKLEDQDKAAE 282

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:55:14 ; Search time 18.4867 Seconds
(without alignments)
399.760 Million cell updates/sec

Title: US-10-674-755-11

Perfect score: 485

Sequence: 1 LKEIDSESDYVKEGFRAP.....KXAELEKTEADLKXAVNEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/prodata/1/iaa/5A COMB.pep.*
- 2: /cgn2_6/prodata/1/iaa/5B COMB.pep.*
- 3: /cgn2_6/prodata/1/iaa/6A COMB.pep.*
- 4: /cgn2_6/prodata/1/iaa/6B COMB.pep.*
- 5: /cgn2_6/prodata/1/iaa/PCITUS COMB.pep.*
- 6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	485	100.0	99	2	US-08-710-749-10
2	485	100.0	99	4	US-09-147-875A-11
3	485	100.0	204	4	US-08-529-055-51
4	485	100.0	8991	4	US-08-714-741-32
5	469	96.7	99	2	US-08-710-749-11
6	469	96.7	198	4	US-08-529-055-61
7	469	96.7	619	1	US-08-465-746-2
8	469	96.7	619	1	US-08-214-164-2
9	469	96.7	619	2	US-08-467-852A-3
10	469	96.7	619	2	US-08-246-636-2
11	469	96.7	619	2	US-08-247-491A-3
12	469	96.7	619	2	US-08-319-795-2
13	469	96.7	619	2	US-08-468-985-2
14	469	96.7	619	3	US-08-312-949-2
15	469	96.7	648	1	US-08-072-070-2
16	469	96.7	648	1	US-08-469-434-2
17	469	96.7	648	1	US-08-214-222-2
18	469	96.7	648	2	US-08-467-852A-2
19	469	96.7	648	2	US-08-468-718-2
20	469	96.7	648	2	US-08-247-491A-2
21	469	96.7	648	3	US-08-446-201-3
22	469	96.7	695	1	US-08-127-499A-23
23	469	96.7	695	1	US-08-482-847-23
24	462.5	95.4	100	4	US-09-147-875A-12
25	457	94.2	288	3	US-08-312-949-4
26	457	94.2	288	3	US-08-446-201-4
27	446.5	92.1	289	1	US-08-072-070-4

28	446.5	92.1	289	1	US-08-469-434-4	Sequence 4, Appli
29	446.5	92.1	289	1	US-08-214-222-4	Sequence 4, Appli
30	446.5	92.1	289	2	US-08-467-852A-5	Sequence 5, Appli
31	446.5	92.1	289	2	US-08-468-718-4	Sequence 4, Appli
32	446.5	92.1	289	2	US-08-247-491A-5	Sequence 5, Appli
33	446	92.0	99	4	US-09-147-875A-13	Sequence 13, Appli
34	442	91.1	99	2	US-08-710-749-12	Sequence 12, Appli
35	442	91.1	195	4	US-08-529-055-71	Sequence 71, Appli
36	423	87.2	623	4	US-08-714-741-47	Sequence 60, Appli
37	414	85.4	170	4	US-08-529-055-60	Sequence 47, Appli
38	414	85.4	181	4	US-08-529-055-42	Sequence 42, Appli
39	414	85.4	864	4	US-08-714-741-40	Sequence 40, Appli
40	413	85.2	99	2	US-08-710-749-17	Sequence 17, Appli
41	412.5	85.1	100	4	US-09-147-875A-10	Sequence 10, Appli
42	411	84.7	99	4	US-09-147-875A-16	Sequence 16, Appli
43	411	84.7	188	4	US-08-529-055-59	Sequence 59, Appli
44	410	84.5	1231	4	US-08-714-741-41	Sequence 41, Appli
45	402	82.9	141	4	US-09-286-981B-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-710-749-10
; Sequence 10, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-10

Query Match 100.0%; Score 485; DB 2; Length 99;
Best Local Similarity 100.0%; Pred. No. 2.1e-37;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LKEIDSESDYVKEGFRAPLOSELDKAKLSKEELSDKIDELDAETAKLEDQKAAE 60
DB 1 LKEIDSESDYVKEGFRAPLOSELDKAKLSKEELSDKIDELDAETAKLEDQKAAE 60

RESULT 3
US-08-529-055-51
Sequence 51, Application US/08529055
Patent No. 6592876
GENERAL INFORMATION:
APPLICANT: Briles, David E.
APPLICANT: McDaniel, Larry S.
APPLICANT: Swiatlo, Edwin
APPLICANT: Yother, Janet
APPLICANT: Brooks-Walter, Alexis
TITLE OF INVENTION: Pneumococcal Genes, Portions
Thereof, Expression Products
TITLE OF INVENTION: Thereof, Expression Products
TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
Portions and Products
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford, P.C.
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/529,055
FILING DATE: 15-SEP-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454312-2400
TELECOMMUNICATION INFORMATION:

/
 / ADDRESS: Curtis, Morris & Safford, P.C.
 / STREET: 530 Fifth Avenue
 / CITY: New York
 / STATE: New York
 / COUNTRY: U.S.
 / ZIP: 10036
 /
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: PatentIn Release #1.0, Version #1.30
 /
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/714,741
 / FILING DATE: 16-SEP-1996
 / CLASSIFICATION: 435
 /
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Frommer Esq., William S.
 / REGISTRATION NUMBER: 25,506
 / REFERENCE/DOCKET NUMBER: 454312-2460
 /
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (212) 840-3333
 / TELEFAX: (212) 840-0712
 /
 / INFORMATION FOR SEQ ID NO: 32:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 8991 amino acids
 / TYPE: amino acid
 / STRANDEDNESS: single
 / TOPOLOGY: linear
 /
 / MOLECULE TYPE: amino acid
 /
 / US-08-714-741-32


```
Query Match      100.0%; Score 485; DB 4; Length 8991;
Best Local Similarity 100.0%; Pred. No. 4.5e-35;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLELSKIDELDAETAKLEDQKAAE 60
DB 5322 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLELSKIDELDAETAKLEDQKAAE 5381
QY 61 ENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 99
DB 5382 ENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 5420

RESULT 5
US-08-710-749-11
; Sequence 11, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-710-749-11

Query Match      96.7%; Score 469; DB 2; Length 99;
Best Local Similarity 96.0%; Pred. No. 6.2e-36;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLELSKIDELDAETAKLEDQKAAE 60
DB 1 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLELSKIDELDAETAKLEDQKAAE 60
QY 61 ENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 99
DB 61 ENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 99

RESULT 6
US-08-529-055-61
; Sequence 61, Application US/08529055
; Patent No. 6592876
```

```
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-61

Query Match      96.7%; Score 469; DB 4; Length 198;
Best Local Similarity 96.0%; Pred. No. 1.4e-35;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLELSKIDELDAETAKLEDQKAAE 60
DB 1 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLELSKIDELDAETAKLEDQKAAE 60
QY 61 ENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 99
DB 61 ENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 99

RESULT 7
US-08-465-746-2
; Sequence 2, Application US/08465746
; Patent No. 5679768
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Yother, Janet L.
; APPLICANT: McDaniel, Larry S.
; TITLE OF INVENTION: EPITOPIC REGIONS OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEIN A
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
```

```
;
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,746
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/048,896
; FILING DATE:
; APPLICATION NUMBER: US 07/656,773
; FILING DATE: 15-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,698
; FILING DATE: 12-FEB-1992
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 521-0378
; TELEX: LUKPAT WASHINGTON
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-465-746-2

Query Match 96.7%; Score 469; DB 1; Length 619;
Best Local Similarity 96.0%; Pred. No. 5.4e-35;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGFRAPLQSELDKAKKLSKLELSKDIDELDAEIAKLEDLQKAAE 60
Db 223 LKEIDSESDYVKEGFRAPLQSKLDKAKKLSKLELSKDIDELDAEIAKLEDLQKAAE 282
QY 61 ENNVEDYFKEGLEKTTIAAKKAELEKTEADLKKA VNEPE 99
Db 283 ENNVEDYFKEGLEKTTIAAKKAELEKTEADLKKA VNEPE 321

RESULT 8
US-08-214-164-2
; Sequence 2, Application US/08214164
; Patent No. 5728387
; GENERAL INFORMATION:
; APPLICANT: BRILES, DAVID E.
; APPLICANT: YOTTER, JANET L.
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Lcd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/214,164
; FILING DATE: 17-MAR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/656,773
; FILING DATE: 15-FEB-1991
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:

; NAME: Berkstresser, Jerry W.
; REGISTRATION NUMBER: 22,651
; REFERENCE/DOCKET NUMBER: 6102-137
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 521-0813
; TELEX: LUKPAT WASHINGTON
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-214-164-2

Query Match 96.7%; Score 469; DB 1; Length 619;
Best Local Similarity 96.0%; Pred. No. 5.4e-35;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGFRAPLQSELDKAKKLSKLELSKDIDELDAEIAKLEDLQKAAE 60
Db 223 LKEIDSESDYVKEGFRAPLQSKLDKAKKLSKLELSKDIDELDAEIAKLEDLQKAAE 282
QY 61 ENNVEDYFKEGLEKTTIAAKKAELEKTEADLKKA VNEPE 99
Db 283 ENNVEDYFKEGLEKTTIAAKKAELEKTEADLKKA VNEPE 321

RESULT 9
US-08-467-852A-3
; Sequence 3, Application US/08467852A
; Patent No. 5856170
; GENERAL INFORMATION:
; APPLICANT: BRILES, David E.
; APPLICANT: YOTTER, Janet L.
; APPLICANT: MCDANIEL, Larry S.
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP
; STREET: 745 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10151
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,852A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOWALSKI, Thomas J.
; REGISTRATION NUMBER: 32,147
; REFERENCE/DOCKET NUMBER: 454312-2064
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-588-0800
; TELEFAX: 212-588-0500
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-467-852A-3

Query Match 96.7%; Score 469; DB 2; Length 619;
Best Local Similarity 96.0%; Pred. No. 5.4e-35;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

Qy	1	LKEID	DES	SED	YK	G	F	R	A	P	O	S	E	L	D	A	K	A	K	S	K	L	E	E	S	D	K	I	D	L	D	A	E	I	A	K	L	E	D	L	K	A	E	60	
Db	223	LKEID	SES	E	D	YK	G	F	R	A	P	O	S	E	L	D	A	K	A	K	S	K	L	E	E	S	D	K	I	D	L	D	A	E	I	A	K	L	E	D	L	K	A	E	282
Qy	61	ENN	N	V	E	D	Y	F	K	E	G	L	E	K	T	I	A	A	K	A	E	L	E	K	T	E	A	D	L	K	A	V	N	P	E	99									
Db	283	ENN	N	V	E	D	Y	F	K	E	G	L	E	K	T	I	A	A	K	A	E	L	E	K	T	E	A	D	L	K	A	V	N	P	E	321									

```

RESULT 10
US-08-246-636-2
; Sequence 2, Application US/08246636
; Patent No. 5965141
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Yother, Janet L.
; APPLICANT: MCDaniel, Larry S
; APPLICANT: Wu, Hong-Yin
; TITLE OF INVENTION: EPITOPIC REGIONS OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEIN A
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
;

```

ADDRESSEE: Shoemaker and Mattare, Ltd
STREET: Suite 1203, 2001 Jefferson Davis Highway
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202-0286

```

? ZIP: 22202-0286
?
? COMPUTER READABLE FORM:
?
? MEDIUM TYPE: Floppy disk
?
? COMPUTER: IBM PC compatible
?
? OPERATING SYSTEM: PC-DOS/MS-DOS
?
? SOFTWARE: PatentIn Release #1.0, Version #1.25
?
? CURRENT APPLICATION DATA:
?
? APPLICATION NUMBER: US/08/246,636
?
? FILING DATE: 20-MAY-1994
?

```

```

:
:
: CLASSIFICATION: 435
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/656,773
: FILING DATE: 15-FEB-1991
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/835,698
: FILING DATE: 12-FEB-1992
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/048,896
:
:
:

```

FILING DATE: 20-APR-1993
TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 415-0810
TELEFAX: (703) 415-0813

TELEX: LUKPAT WASHINGTON
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 619 amino acids

```
;
;
; TYPE: amino acid
; TOPOLOGY: linear
```

US-08-246-636-2

Query Match 96.7%: S

Best Local Similarity 96.0%;
Matches 95; Conservative 3;

QY 1 LKEIDSESDYVKEGFRAPL

223 LKEIDSESEDYAKEGFRAPL

QY 61 ENNNVEDYFKEGLEKTI AAKK

283 ENNVEDYFKEGLEKTI AAKK

.....

RESULT 11

US-08-247-491A-3
; Sequence 3, Application US/08247491A
; Patent No. 5965400
; GENERAL INFORMATION:

APPLICANT: BRILES, David E.
APPLICANT: YOTHER, Janet L.
TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: FROMMER LAWRENCE & HAUG LLP
STREET: 745 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA

```

? ZIP: 10151
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/247,491A
? FILING DATE: 23-JUN-1994
?
```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KOWALSKI, Thomas J.
REGISTRATION NUMBER: 32,147
REFERENCE/DOCKET NUMBER: 454313-2041
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-588-0800
TELEFAX: 212-588-0500

```

, INFORMATION FOR SEQ ID NO: 3:
,
, SEQUENCE CHARACTERISTICS:
, LENGTH: 619 amino acids
, TYPE: amino acid
, STRANDEDNESS: n/a
, TOPOLOGY: linear
,
, MOLECULE TYPE: amino acid
US-08-247-491A-3

```

```

Query Match      96.7%; Score 469; DB 2; Length 619;
Best Local Similarity 96.0%; Pred. No. 5.4e-35;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKIDISDSBDYKGFAPLQSELDAKQKLSKLSLSLSLSKIDELDAEIAKLEDLQKAAE 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 223 LKEIDSESDYAKGFAPLQSKLDKAKKLSKLSLSLSKIDELDAEIAKLEDLQKAAE 282

Qy 61 ENNVEDYFKGGLKTIKAAKAELEKTEADLKVAVNEPE 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 283 ENNVEDYFKGGLKTIKAAKAELEKTEADLKVAVNEPE 321
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

RESULT 12
US-08-319-795-2
; Sequence 2, Application US/08319795
. Patent NO. 5980909

? PATENT NO. 3260393
 ? GENERAL INFORMATION:
 ? APPLICANT: Briles, David E.
 ? APPLICANT: Yother, Janet L.
 ? APPLICANT: McDaniel, Larry S
 ? TITLE OF INVENTION: Epitopic Regions of Pneumococcal Surface
 ? NUMBER OF SEQUENCES: 20
 ? CORRESPONDENCE ADDRESS:
 ? ADDRESSEE: Sheomaker and Mattare, Ltd.
 ? STREET: 1203 Crystal Plaza Bldg. 1, 2001 Jefferson
 ? STREET: Davis Highway
 ? CITY: Arlington
 ? STATE: Virginia
 ? COUNTRY: U.S.A.
 ? ZIP: 22202-0286
 ? COMPUTER READABLE FORM:

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/319,795
; FILING DATE:
; APPLICATION NUMBER: US/08/319,795
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/246,636
; FILING DATE: 20-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/048,896
; FILING DATE: 20-APR-1993
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-319-795-2

Query Match          96.7%; Score 469; DB 2; Length 619;
Best Local Similarity 96.0%; Pred. No. 5.4e-35;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSDKIDELDAEIAKLEDLKAAE 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 223 LKEIDSESEDYAKGFRAPLQSKLDKAKKLSKLELSDKIDELDAEIAKLEDLKAAE 282

Qy 61 ENNVEDYFKEGLEKTIKAAKAELEKTEADLKKA VNEPE 99
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 283 ENNVEDYFKEGLEKTIKAAKAELEKTEADLKKA VNEPE 321

RESULT 13
US-08-468-985-2
; Sequence 2, Application US/08468985
; Patent No. 5997882
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Yother, Janet L.
; APPLICANT: McDaniel, Larry S
; TITLE OF INVENTION: Epitopic Regions of Pneumococcal Surface
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheomaker and Mattare, Ltd.
; STREET: 1203 Crystal Plaza Bldg. 1, 2001 Jefferson
; STREET: Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,985
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,698
; FILING DATE:
; APPLICATION NUMBER: US 07/835,698
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/246,636
; FILING DATE: 20-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/048,896
; FILING DATE: 20-APR-1993
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-985-2

Query Match          96.7%; Score 469; DB 2; Length 619;
Best Local Similarity 96.0%; Pred. No. 5.4e-35;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSDKIDELDAEIAKLEDLKAAE 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 223 LKEIDSESEDYAKGFRAPLQSKLDKAKKLSKLELSDKIDELDAEIAKLEDLKAAE 282

Qy 61 ENNVEDYFKEGLEKTIKAAKAELEKTEADLKKA VNEPE 99
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 283 ENNVEDYFKEGLEKTIKAAKAELEKTEADLKKA VNEPE 321

RESULT 14
US-08-312-949-2
; Sequence 2, Application US/08312949
; Patent No. 6027734
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Wu, Hong-Yin
; TITLE OF INVENTION: MUCOSAL ADMINISTRATION OF
; TITLE OF INVENTION: PNEUMOCOCCAL ANTIGENS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,949
; FILING DATE: 30-SEP-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2049
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-312-949-2

Query Match 96.7%; Score 469; DB 3; Length 619;
Best Local Similarity 96.0%; Pred. No. 5.4e-35;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGFRAPLQSELDKAKSKLELSKIDELDAEIAKLEDLQKAAE 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 223 LKEIDSESDYAKGFRAPLQSKLDAKAKSKLELSKIDELDAEIAKLEDLQKAAE 282
QY 61 ENNVEDYFKGLEKTIAAKKAELEKTEADLKKA VNEPE 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 283 ENNVEDYFKGLEKTIAAKKAELEKTEADLKKA VNEPE 321

RESULT 15
US-08-072-070-2
; Sequence 2, Application US/08072070
; Patent No: 5476929
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: Yother, Janet L
; APPLICANT: McDaniel, Larry S
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/072,070
; FILING DATE: 19930603
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/835,698
; FILING DATE: 12-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/656,773
; FILING DATE: 15-FEB-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 521-0378
; TELEX: LUKPAT WASHINGTON
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 648 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-072-070-2

Query Match 96.7%; Score 469; DB 1; Length 648;
Best Local Similarity 96.0%; Pred. No. 5.7e-35;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGFRAPLQSELDKAKSKLELSKIDELDAEIAKLEDLQKAAE 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 223 LKEIDSESDYAKGFRAPLQSKLDAKAKSKLELSKIDELDAEIAKLEDLQKAAE 282
QY 61 ENNVEDYFKGLEKTIAAKKAELEKTEADLKKA VNEPE 99

DB 283 ENNVEDYFKGLEKTIAAKKAELEKTEADLKKA VNEPE 321
Search completed: June 21, 2005, 10:25:19
Job time : 19.4867 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.
OM protein - protein search, using sw model
Run on: June 21, 2005, 10:12:15 ; Search time 63.2388 Seconds
(without alignments)
601.118 Million cell updates/sec

Title: US-10-674-755-11
Perfect score: 485
Sequence: 1 LKEIDSESDYVKEGFRAP.....KKAELEKTEADLKAVNEPE 99

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	485	100.0	99	15	US-10-674-755-11
2	485	100.0	204	15	US-10-299-636-66
3	469	96.7	138	15	US-10-299-636-76
4	469	96.7	354	15	US-10-299-636-105
5	469	96.7	588	10	US-10-299-636-96
6	469	96.7	619	10	US-09-882-774-1
7	469	96.7	619	15	US-10-282-122A-73702
8	469	96.7	619	16	US-10-414-532-72
9	462.5	95.4	100	15	US-10-674-755-12
10	446	92.0	99	15	US-10-674-755-13
11	442	91.1	195	15	US-10-299-636-86
					Sequence 11, Appl
					Sequence 66, Appl
					Sequence 76, Appl
					Sequence 105, Appl
					Sequence 96, Appl
					Sequence 1, Appl
					Sequence 73702, A
					Sequence 72, Appl
					Sequence 12, Appl
					Sequence 13, Appl
					Sequence 86, Appl

12	414	85.4	170	15	US-10-299-636-75	Sequence 75, Appl
13	414	85.4	181	15	US-10-299-636-57	Sequence 57, Appl
14	414	85.4	643	15	US-10-299-636-95	Sequence 95, Appl
15	414	85.4	670	9	US-09-748-875-63	Sequence 63, Appl
16	414	85.4	670	10	US-09-298-523B-63	Sequence 63, Appl
17	414	85.4	690	9	US-09-748-875-61	Sequence 61, Appl
18	414	85.4	690	10	US-09-298-523B-61	Sequence 61, Appl
19	414	85.4	691	9	US-09-748-875-1	Sequence 1, Appl
20	414	85.4	691	10	US-09-298-523B-1	Sequence 1, Appl
21	414	85.4	701	9	US-09-748-875-62	Sequence 62, Appl
22	414	85.4	701	10	US-09-298-523B-62	Sequence 62, Appl
23	414	85.4	707	9	US-09-748-875-2	Sequence 2, Appl
24	414	85.4	707	10	US-09-298-523B-2	Sequence 2, Appl
25	414	85.4	711	9	US-09-748-875-3	Sequence 3, Appl
26	414	85.4	711	10	US-09-298-523B-3	Sequence 3, Appl
27	414	85.4	739	17	US-10-732-923-3294	Sequence 3294, Ap
28	414	85.4	929	9	US-09-748-875-60	Sequence 60, Appl
29	414	85.4	929	10	US-09-298-523B-60	Sequence 60, Appl
30	414	85.4	929	15	US-10-299-636-94	Sequence 94, Appl
31	412.5	85.1	100	15	US-10-674-755-10	Sequence 10, Appl
32	411	84.7	99	15	US-10-674-755-16	Sequence 16, Appl
33	411	84.7	188	15	US-10-299-636-74	Sequence 74, Appl
34	402	82.9	141	14	US-10-254-995-2	Sequence 2, Appl
35	402	82.9	589	9	US-09-748-875-14	Sequence 14, Appl
36	402	82.9	589	10	US-09-298-523B-14	Sequence 14, Appl
37	402	82.9	589	15	US-10-299-636-97	Sequence 97, Appl
38	400	82.5	204	15	US-10-299-636-73	Sequence 73, Appl
39	393	81.0	99	15	US-10-674-755-15	Sequence 15, Appl
40	389	80.2	206	15	US-10-299-636-69	Sequence 69, Appl
41	386	79.6	99	15	US-10-674-755-14	Sequence 14, Appl
42	352.5	72.7	100	15	US-10-674-755-2	Sequence 2, Appl
43	348.5	71.9	100	15	US-10-674-755-3	Sequence 3, Appl
44	343.5	70.8	194	15	US-10-299-636-79	Sequence 79, Appl
45	339.5	70.0	98	15	US-10-674-755-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-10-674-755-11
; Sequence 11, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-11

Query Match 100.0%; Score 485; DB 15; Length 99;
Best Local Similarity 100.0%; Pred. No. 3.1e-32;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEDQKAAE	60
Db	1	LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEDQKAAE	60
QY	61	ENNVNEDYFKEGLEKTIAAKKALEKTEADLKAVNEPE	99
Db	61	ENNVNEDYFKEGLEKTIAAKKALEKTEADLKAVNEPE	99

RESULT 2

[illegible]


```
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 96
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-96

Query Match          96.7%; Score 469; DB 15; Length 588;
Best Local Similarity 96.0%; Pred. No. 4.9e-30;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGFAPLQSELDKQAKLSKLEELSDKIDELDAEIAKLSDQLKAAE 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 192 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAEIAKLSDQLKAAE 251
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 61 ENNVEDYFKEGLEKTTIAAKKAELEKTEADLKKAVNEPE 99
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 252 ENNVEDYFKEGLEKTTIAAKKAELEKTEADLKKAVNEPE 290
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 6
US-09-882-774-1
; Sequence 1, Application US/09882774
; Publication No. US20030021795A1
; GENERAL INFORMATION:
; APPLICANT: Houston, Robert
; APPLICANT: Hodges, Robert
; TITLE OF INVENTION: Use of Coiled-Coil Structural Scaffold to Generate
; TITLE OF INVENTION: Structure-Specific Peptides
; FILE REFERENCE: 003592-007
; CURRENT APPLICATION NUMBER: US/09/882,774
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,892
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/213,387
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-882-774-1

Query Match          96.7%; Score 469; DB 10; Length 619;
Best Local Similarity 96.0%; Pred. No. 4.9e-30;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGFAPLQSELDKQAKLSKLEELSDKIDELDAEIAKLSDQLKAAE 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 223 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAEIAKLSDQLKAAE 282
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 61 ENNVEDYFKEGLEKTTIAAKKAELEKTEADLKKAVNEPE 99
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 283 ENNVEDYFKEGLEKTTIAAKKAELEKTEADLKKAVNEPE 321
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 7
US-10-282-122A-73702
; Sequence 73702, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert

Query Match          96.7%; Score 469; DB 16; Length 619;
Best Local Similarity 96.0%; Pred. No. 4.9e-30;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73702
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-282-122A-73702

Query Match          96.7%; Score 469; DB 15; Length 619;
Best Local Similarity 96.0%; Pred. No. 4.9e-30;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGFAPLQSELDKQAKLSKLEELSDKIDELDAEIAKLSDQLKAAE 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 223 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAEIAKLSDQLKAAE 282
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 61 ENNVEDYFKEGLEKTTIAAKKAELEKTEADLKKAVNEPE 99
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 283 ENNVEDYFKEGLEKTTIAAKKAELEKTEADLKKAVNEPE 321
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 8
US-10-414-532-72
; Sequence 72, Application US/10414532
; Publication No. US20040101531A1
; GENERAL INFORMATION:
; APPLICANT: CURTISS III, ROY
; APPLICANT: KANG, HO YOUNG
; TITLE OF INVENTION: IMPROVED IMMUNOGENIC COMPOSITIONS AND VACCINES COMPRISING
; FILE REFERENCE: 56029-40437
; CURRENT APPLICATION NUMBER: US/10/414,532
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: 60/372,710
; PRIOR FILING DATE: 2002-04-16
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 72
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-414-532-72

Query Match          96.7%; Score 469; DB 16; Length 619;
Best Local Similarity 96.0%; Pred. No. 4.9e-30;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 LKEIDSDSDYKVEGFRAPLQSLDAKQAKLSKLELSKIDELDAEIAKLEDLQKAAE 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 223 LKEIDSESEDYAKVEGFRAPLQSKLDAKAKLSKLELSKIDELDAEIAKLEDLQKAAE 282

Qy 61 ENNVEDYFKGLEKTTIAAKAELEKTEADLKKA VNEPE 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 283 ENNVEDYFKGLEKTTIAAKAELEKTEADLKKA VNEPE 321

RESULT 9
US-10-674-755-12
; Sequence 12, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-12

Query Match 95.4%; Score 462.5; DB 15; Length 100;
Best Local Similarity 96.0%; Pred. No. 2.1e-30;
Matches 96; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 1 LKEIDSDSDYKVEGFRAPLQSLDAKQAKLSKLELSKIDELDAEIAKLEDLQ-KAA 59
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 LKEIDSESEDYAKVEGFRAPLQSKLDAKQAKLSKLELSKIDELDAEIAKLEDLQKAA 60

Qy 60 ENNVEDYFKGLEKTTIAAKAELEKTEADLKKA VNEPE 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 ENNVEDYFKGLEKTTIAAKAELEKTEADLKKA VNEPE 100

RESULT 10
US-10-674-755-13
; Sequence 13, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-13

Query Match 92.0%; Score 446; DB 15; Length 99;
Best Local Similarity 91.9%; Pred. No. 4.6e-29;
Matches 91; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYKVEGFRAPLQSLDAKQAKLSKLELSKIDELDAEIAKLEDLQKAAE 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 LKEIDSESEDYAKVEGFRAPLHSHKLDKAKLSKLELSKIDELDAEIAKLEDLQKAVE 60

Qy 61 ENNVEDYFKGLEKTTIAAKAELEKTEADLKKA VNEPE 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```
Db 61 ENNVEDYSTEGLEKTTIAAKKTELEKTEADLKKA VNEPE 99

RESULT 11
US-10-299-636-86
; Sequence 86, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-86

Query Match 91.1%; Score 442; DB 15; Length 195;
Best Local Similarity 90.9%; Pred. No. 2.1e-28;
Matches 90; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYKVEGFRAPLQSLDAKQAKLSKLELSKIDELDAEIAKLEDLQKAAE 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 LKEIDSESEDYAKVEGFRAPLHSHKLDKAKKLSKLELSKIDELDAEIAKLEDLQKAVE 60

Qy 61 ENNVEDYFKGLEKTTIAAKAELEKTEADLKKA VNEPE 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 ENNVEDYSTEGLEKTTIAAKKTELEKTEADLKKA VNEPE 99

RESULT 12
US-10-299-636-75
; Sequence 75, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-75
```

Query Match 85.4%; Score 414; DB 15; Length 170;
Best Local Similarity 86.9%; Pred. No. 3.4e-26;
Matches 86; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
QY 1 LKEIDSDSDYKGEGRAPLQSELDKQAKLSKLELSDKIDELDAETAKLEDQKAAE 60
DB 1 LKEIDSDSDYKGEGRAPLQSELDKQAKLSKLELSDKIDELDAETAKLEDQKAAE 60
QY 61 ENNVEDYFKEGLEKTTAAKAELEKTEADLKAVNEPE 99
DB 61 GNNVEAYFKEGLEKTTAEKAELEKAEADLKAVDEPE 99

RESULT 13
US-10-299-636-57
; Sequence 57, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 57
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-57

Query Match 85.4%; Score 414; DB 15; Length 181;
Best Local Similarity 86.9%; Pred. No. 3.7e-26;
Matches 86; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
QY 1 LKEIDSDSDYKGEGRAPLQSELDKQAKLSKLELSDKIDELDAETAKLEDQKAAE 60
DB 1 LKEIDSDSDYKGEGRAPLQSELDKQAKLSKLELSDKIDELDAETAKLEDQKAAE 60
QY 61 ENNVEDYFKEGLEKTTAAKAELEKTEADLKAVNEPE 99
DB 61 GNNVEAYFKEGLEKTTAEKAELEKAEADLKAVDEPE 99

RESULT 14
US-10-299-636-95
; Sequence 95, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19

; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 95
; LENGTH: 643
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-95

Query Match 85.4%; Score 414; DB 15; Length 643;
Best Local Similarity 86.9%; Pred. No. 1.5e-25;
Matches 86; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
QY 1 LKEIDSDSDYKGEGRAPLQSELDKQAKLSKLELSDKIDELDAETAKLEDQKAAE 60
DB 245 LKEIDSDSDYKGEGRAPLQSELDKQAKLSKLELSDKIDELDAETAKLEDQKAAE 304
QY 61 ENNVEDYFKEGLEKTTAAKAELEKTEADLKAVNEPE 99
DB 305 GNNVEAYFKEGLEKTTAEKAELEKAEADLKAVDEPE 343

RESULT 15
US-09-748-875-63
; Sequence 63, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: Briles et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 63
; LENGTH: 670
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-63

Query Match 85.4%; Score 414; DB 9; Length 670;
Best Local Similarity 86.9%; Pred. No. 1.6e-25;
Matches 86; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
QY 1 LKEIDSDSDYKGEGRAPLQSELDKQAKLSKLELSDKIDELDAETAKLEDQKAAE 60
DB 498 LKEIDSDSDYKGEGRAPLQSELDKQAKLSKLELSDKIDELDAETAKLEDQKAAE 557
QY 61 ENNVEDYFKEGLEKTTAAKAELEKTEADLKAVNEPE 99
DB 558 GNNVEAYFKEGLEKTTAEKAELEKAEADLKAVDEPE 596

Search completed: June 21, 2005, 11:18:33
Job time : 64.2388 secs

This Page Blank (uspto)

J. Bacteriol. 174: 601-609, 1992
A:Title: Structural properties and evolutionary relationships of PspA, a surface protein
A:Reference number: A41971; MUID: 92105030; PMID: 1729249
A:Accession: A41971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-619 <YOT>
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:M74122; NID:g153840; PIDN:AAA2701
A:Note: sequence extracted from NCBI backbone (NCBIN:75635, NCBIPI:75636)
R:Talkington, D.F.; Crimmins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.

```
Infect. Immun. 59, 1285-1289, 1991
A;Title: A 43-kilodalton pneumococcal surface protein, PspA: isolation, protective ability
A;Reference number: A60282; MUID:91169598; PMID:2004810
A;Accession: A60282
A;Molecule type: protein
A;Residues: 32-76 <TAL>
A;Experimental source: strain JY2008
C;Genetics:
A;Gene: pspA
F;1-31/Domain: signal sequence #status predicted <SIG>
F;32-619/Product: surface protein pspA #status predicted <MAT>
F;411-430/Domain: cpl repeat homology <CP01>
F;431-450/Domain: cpl repeat homology <CP02>
F;451-470/Domain: cpl repeat homology <CP03>
F;471-490/Domain: cpl repeat homology <CP04>
F;491-510/Domain: cpl repeat homology <CP05>
F;511-530/Domain: cpl repeat homology <CP06>
F;531-550/Domain: cpl repeat homology <CP07>
F;551-570/Domain: cpl repeat homology <CP08>
F;571-591/Domain: cpl repeat homology <CP09>
F;592-611/Domain: cpl repeat homology <CP10>

Query Match 96.7%; Score 469; DB 2; Length 619;
Best Local Similarity 96.0%; Pred. No. 3.3e-24;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEDSDSDYVKEGFRAPLOSELDAKQAKLSKLELSKIDELDAEIAKLELDQKAAE 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 223 LKEDSESDYAKGFRAPLOSKLDKAKKLSKLELSKIDELDAEIAKLELDQKAAE 282
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 61 ENNVEDYFKEGLEKTTIAAKKAELEKTEADLKAVNEPE 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 283 ENNVEDYFKEGLEKTTIAAKKAELEKTEADLKAVNEPE 321
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 3
F95013
pneumococcal surface protein A [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: F95013
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid-
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzap-
neon, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: F95013
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-744 <KUR>
A;Cross-references: UNIPROT:Q97T39; GB:AE005672; PIDN:AAK74303.1; PID:g14971584; GSPDB:G
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP0117

Query Match 29.0%; Score 140.5; DB 2; Length 744;
Best Local Similarity 35.6%; Pred. No. 0.027;
Matches 47; Conservative 14; Mismatches 30; Indels 41; Gaps 5;

Qy 2 KEIDE-----SDSDYVKEGFRAPLOSELDAKQAKLSK-----LELSDKI-----D 43
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 314 KEISNLEILGGADPEDDT-----AALQNKLAAKKAELEKQTELEKLDLSLQPEGKTQD 368
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 44 ELDAEIAKLELDQKAAENNNVEDYFKE-----GLEKTTIAAKKAELE 85
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 369 ELQKAEAEALDKADELQNKVADLEKEISNLEILGGADSDDTAALQNKLATKKAELE 428
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 86 KTEADLKAVNEPE 97
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 429 KTKELDAALNE 440
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 4

```
T30845
probable DNA repair protein RAD50 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C;Accession: T30845
R;Kim, K.K.; Daud, A.I.; Wong, S.C.; Pajak, L.; Tsai, S.C.; Wang, H.; Henzel, W.J.; Field
J. Biol. Chem. 271, 29255-29264, 1996
A;Title: Mouse RAD50 has limited epitopic homology to p53 and is expressed in the adult
A;Reference number: Z20899; MUID:97067183; PMID:8910585
A;Accession: T30845
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1312 <KIM>
A;Cross-references: UNIPROT:P70388; EMBL:U66887; NID:g1575574; PID:g1575575; PIDN:AAAC528;
C;Genetics:
A;Gene: RAD50
A;Map position: 11
C;Superfamily: RAD50 protein
C;Keywords: DNA repair

Query Match 23.4%; Score 113.5; DB 2; Length 1312;
Best Local Similarity 35.5%; Pred. No. 3;
Matches 27; Conservative 19; Mismatches 21; Indels 9; Gaps 1;

Qy 22 QSELDAKQAKLSKLELSKIDELDAEIAKLELDQKAAENNNVEDYFKEGLEKTTIAAKK 81
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 454 QSELHVRSELQQLQEGSSDRILELDQELTKAERLSKAEKNSSIE-----TLKAEV 504
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 82 AELEKTEADLKAVNEPE 97
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 505 MSLQNEKADLDRLSRK 520
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 5
S48396
tropomyosin TPM2 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YII138c
C;Species: Saccharomyces cerevisiae
C;Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004
C;Accession: S48396; A56490
R;Churcher, C.
submitted to the EMBL Data Library, September 1994
A;Reference number: S48310
A;Accession: S48396
A;Molecule type: DNA
A;Residues: 1-161 <CHU>
A;Cross-references: UNIPROT:P40414; GB:Z47047; EMBL:Z38059; NID:g603997; PID:g763208; MUI
R;Drees, B.; Brown, C.; Barrell, B.G.; Bretscher, A.
J. Cell Biol. 128, 383-392, 1995
A;Title: Tropomyosin is essential in yeast, yet the TPM1 and TPM2 products perform disti
A;Reference number: A56490; MUID:95146545; PMID:7844152
A;Accession: A56490
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-161 <DRE>
A;Cross-references: GB:Z47047; GB:Z38059; NID:g603997; PID:g763208
C;Genetics:
A;Gene: SGD:TPM2
A;Cross-references: SGD:S0001400; MIPS:YII138c
A;Map position: 9L
C;Superfamily: tropomyosin TPM1
C;Keywords: cytoskeleton

Query Match 23.0%; Score 111.5; DB 2; Length 161;
Best Local Similarity 33.3%; Pred. No. 0.5;
Matches 35; Conservative 24; Mismatches 31; Indels 15; Gaps 4;

Qy 1 LKEDSDSDYVKEGFRAPLOSELDAKQAKLSKLELSKIDELDAEIAKLELDQ----- 56
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 11 LKLESESWQEKY--EELRQLKELEQSNTEKENEIKLSAKNEQLDSEVEKLSQLSDTK 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

QY 57 KAAEENNV-----EDYFKEGLEKTIAAKAELEKTEADLKAAVNE 97
 Db 69 QLAEDSNLNRNNENYTKNQD-----LRQULEDSEAKLKEAMDK 108

RESULT 6

T34418
 hypothetical protein F12F3.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T34418
 R:Pulton, B.; Wohldmann, P.
 submitted to the EMBL Data Library, July 1998
 A:Description: The sequence of C. elegans cosmid F12F3.
 A:Reference number: Z21521
 A:Accession: T34418
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-3488 <FUL>
 A:Cross-references: EMBL:U80022; PIDN:AAC25885.1; GSPDB:GN00023; CESP:F12F3.3
 A:Experimental source: strain Bristol N2; clone F12F3
 C:Genetics:
 A:Gene: CESP:F12F3.3
 A:Map position: 5
 A:Introns: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1

Query Match 22.8%; Score 110.5; DB 2; Length 3488;
 Best Local Similarity 34.4%; Pred. No. 12;
 Matches 43; Conservative 21; Mismatches 26; Indels 35; Gaps 8;
 QY 2 KEIDES---DSEDYVKEGFRAPLQSELDAAKQAKLSKL-----EELSDKIDELDARI 49
 Db 1009 KETDEKLDAEIAATAKTQEADEKSLDA-QEKIKKVSDDAARKEKELNDKL-KLESRI 1066
 QY 50 A-----KLEDO-----LKAEEENNVEDYF---EGLEKTIAAKAELEKTEA 89
 Db 1067 ATKASADKLEEQAAKAAEVAEAKKQKDEQLKLDTEAAKAAAEKLELEK-QA 1125
 QY 90 DLKKA 94
 Db 1126 QIKKA 1130

RESULT 7

A28313
 glued protein - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
 C:Accession: A28313
 R:Swaroop, A.; Swaroop, M.; Garen, A.
 Proc. Natl. Acad. Sci. U.S.A. 84, 6501-6505, 1987
 A:Title: Sequence analysis of the complete cDNA and encoded polypeptide for the glued ge
 A:Reference number: A28313; MUID:87317680; PMID:2819881
 A:Accession: A28313
 A:Molecule type: DNA; mRNA
 A:Residues: 1-1319 <SWA>
 A:Cross-references: UNIPROT:PI3496
 A>Note: the authors' translation is inconsistent with the nucleotide sequence in the reg
 C:Genetics:
 A:Gene: FlyBase:Gl
 A:Cross-references: FlyBase:FBgn0001108
 A:Introns: 18/2; 479/3
 C:Keywords: cytoskeleton; glycoprotein
 P:397,590,771,889,980,1110,1127,1133,1142/Binding site: carbohydrate (Asn) (covalent) #e

Query Match 22.2%; Score 107.5; DB 2; Length 1319;
 Best Local Similarity 32.5%; Pred. No. 7.4;
 Matches 37; Conservative 24; Mismatches 30; Indels 23; Gaps 5;
 QY 1 LKEIDSDSEDYVKEGFRAPLQSELDAAKQAKLSKL-----EELSDKIDELDAAETAKLEDOL 56
 Db 429 LRDLSDHKHDQK-----LSKELEMKRSEVTELETKESAKIDELEAIVADLQEQV 482

QY 57 KA-----ABENNVEDYFKEGLEKTIAAKAELEKTEADLKAAVNEPE 99
 Db 483 DAALGAEEHVEQLAEKQMELEDKVL-LEEEIAQLEA-LEEVHEQLVESNHSLE 534

RESULT 8

S70531
 bmk2.11 protein precursor - Lyme disease spirochete
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
 C:Accession: S70531
 R:Akins, D.R.; Porcella, S.F.; Popova, T.G.; Shevchenko, D.; Baker, S.I.; Li, M.; Norgard
 Mol. Microbiol. 18, 507-520, 1995
 A:Title: Evidence for in vivo but not in vitro expression of a Borrelia burgdorferi outer
 A:Reference number: S70531; MUID:96342380; PMID:8748034
 A:Accession: S70531
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-233 <AKI>
 A:Cross-references: UNIPROT:Q44739; EMBL:U30617; NID:g3309515; PIDN:AAC46421.1; PID:g119:
 C:Superfamily: outer surface protein F ospF
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-233/Product: bmk2.11 protein #status predicted <MAT>

Query Match 22.1%; Score 107; DB 2; Length 233;
 Best Local Similarity 29.8%; Pred. No. 1.4;
 Matches 37; Conservative 25; Mismatches 34; Indels 28; Gaps 6;
 QY 1 LKEIDES--DSEDYK-----EGFRAPLQ---SELDAAKQAK--LSKLEELSDKI 42
 Db 29 LKNSQNLESSEQNKKTEQETIKQVEGFLEITKDLKDEKTKETKQIQELNKKI 88
 QY 43 DELDAEIAKLE-----DOLKAAEENNVEDYFKEGLEKTIAAKAELEKTEADLKK 93
 Db 89 EKLSKTKTSIETYSEVEEKINKIKLKGKLEDFKE-LEESLAKKKGERKKALQEAQ 147
 QY 94 AVNE 97
 Db 148 KPEE 151

RESULT 9

H69378
 conserved hypothetical protein AF1032 - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C:Accession: H69378
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.N.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
 A:Reference number: A69250; MUID:98049343; PMID:9389475
 A:Accession: H69378
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-886 <KLE>
 A:Cross-references: UNIPROT:O29230; GB:AE001032; GB:AE000782; NID:g2689355; PIDN:AAB90211
 C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 22.1%; Score 107; DB 2; Length 886;
 Best Local Similarity 25.9%; Pred. No. 5.4;
 Matches 35; Conservative 25; Mismatches 29; Indels 46; Gaps 5;
 QY 1 LKEIDESSEDYVKEGFRAPLQSELDAAKQAKLSKLEELSDKIDELDAEI-----49
 Db 303 LRDEVKREG-DLTREA--AGIQALKKABEDNSKLEETIKRIEELERELERFEKSHRLLE 359
 QY 50 -----AKLEDO-----LKAEEENNVEDYFKEGLEKTIAAKKA 82
 Db 360 TLKPKMDRMQGIKAKLEKNLTPDKVKRMVYLLLSNAKEBEKITEKLLK-----LIAKKS 414


```

Query Match          21.2%; Score 103; DB 2; Length 3461;
Best Local Similarity 32.6%; Pred. No. 39;
Matches 43; Conservative 12; Mismatches 35; Indels 42; Gaps 5;

QY      1 LKEIDESDSYVYKGFAPLQSELDAAKQAKSLKLE-----ELSDKI 42
Db      802 LKGILEATELEVDATNRAIOEQLETTOKKADELERKTIENVKKAALNAQNEGLELEKKL 861

QY      43 DEL-----DAEIAKLEDDOLKAAEE--NNNVEDYKPEGLEKTIATAKKAE 83
Db      862 DELIGTVNSAENELELAAPIAEASLSKLDELKRAELFQKLIEN---EGDVSIRAKVAE 918

QY      84  --LEKTEADLKK 93
Db      919 ELKKKPDRELKK 930

RESULT 14
S43074
epidermal growth factor receptor substrate - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S43074; I38525
R:Bernard, O.A.; Mauchauffe, M.; Mecucci, C.; van den Berghe, H.; Berger, R.
Oncogene 9, 1039-1045, 1994
A:Title: A novel gene, AP-1p, fused to HRX in t(1;11)(p32;q23), is not related to AP-4,
A:Reference number: S43074; MUID:94181254; PMID:8134107
A:Accession: S43074
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-896 <BER>
A:Cross-references: UNIPROT:P42566; EMBL:Z29064; NID:g470034; PIDN:CAA82305.1; PID:g4700
R:Wong, W.T.; Kraus, M.H.; Carlomagno, F.; Zelano, A.; Druck, T.; Croce, C.M.; Huebner,
Oncogene 9, 1591-1597, 1994
A:Title: The human eps15 gene, encoding a tyrosine kinase substrate, is conserved in eve
A:Reference number: I38525; MUID:94239734; PMID:8483552
A:Accession: I38525
A>Status: preliminary; translated from GE/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-821, 'M', 823-896 <RES>
A:Cross-references: EMBL:U07707; NID:g466259; PIDN:AAA52101.1; PID:g466260
C:Genetics:
A:Gene: GDB:EP515; AF-1P; MLLT5
A:Cross-references: GDB:360337; OMIM:600051
A:Map position: lp32-lp32

```

```

Query Match      21.0%; Score 102; DB 2; Length 896;
Best Local Similarity 27.7%; Pred. No. 12;
Matches 28; Conservative 23; Mismatches 44; Indels 6; Gaps 2;

Qy 3 EIDESSEDVYKGF--APQSELDAAQAKSL-----EELSDKIDELDAETAKLEDQL 56
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 353 EQDLKEKEDTIKORTSEVDQLQDEVQRENTNLQLOAQKQVQVELLDDELDEQAQLEEQ 412
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

Qy 57 KAAEENNNVEDYFKEGLEKTIAAKKAELEKTEADLLKAVNE 97
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

Db 413 KEVRKKCAEBAQLISSLKAEILTQSOEISITYSYBELAKAREE 453

```

RESULT 15
D72230
conserved hypothetical protein - Thermotoga maritima (strain MSB8)
C/Species: Thermotoga maritima
C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C/Accession: D72230
R/Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.N.
Nature 399, 323-329, 1999
A/Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A/Reference number: A72200; MUID:99287316; PMID:10360571
A/Accession: D72230
A/Status: preliminary

```

A:Molecule type: DNA
A:Residues: 1-852 <ARN>
A:Cross-references: UNIPROT:Q9X1X1; GB:AE001806; GB:AE000512; NID:G4982196; PID:
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM1636
C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match      20.9%; Score 101.5; DB 2; Length 852;
Best Local Similarity 29.2%; Pred. No. 12;
Matches 26; Conservative 23; Mismatches 29; Indels 11; Gaps 3

Qy      6  ESDSEYVVEGFRAPQSELDAAQAKLSLEE----LSDKIDELDARIATKLELDQAKAAEE 61
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      506 EKTEELHRLIGSYEDLEELKDEKKRLKRTKEERHSISQKITAADVQISQENQLK--EI 563

Qy      62  NNNVEDYFKGLEKTTAAKAAELEKTEAD 90
      : | | : | | : | | : | | : | |
Db      564 KGIH-----EAKRETLKEQREEMDLQKSD 587

Search completed: June 21, 2005, 10:11:57
Job time : 11.9 secs

```

Search completed: June 21, 2005, 10:11:57
Job time : 11.9 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:54:24 ; Search time 61.3194 Seconds
(without alignments)
826.751 Million cell updates/sec

Title: US-10-674-755-11

Perfect score: 485

Sequence: 1 LKEIDSESDYVKGFRAP.....KKAELEKTEADLKAVNEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	485	100.0	417	2 Q9LAY3	Q9lay3 streptococc
2	459	96.7	619	2 Q54972	Q54972 streptococc
3	469	96.7	619	2 Q8DR10	Q8dr10 streptococc
4	447	92.2	415	2 Q9LAY1	Q9lay1 streptococc
5	414	85.4	739	2 Q9ROT4	Q9rot4 streptococc
6	414	85.4	820	2 Q9QRT1	Q9rt1 streptococc
7	414	85.4	929	2 Q9KK19	Q9kk19 streptococc
8	414	85.4	929	2 Q9ZAY5	Q9zay5 streptococc
9	405	83.5	99	2 Q8KQK4	Q8kk4 streptococc
10	405	83.5	249	2 Q9L575	Q9l575 streptococc
11	398	82.1	224	2 Q8GNS8	Q8gns8 streptococc
12	398	82.1	437	2 Q9LAY4	Q9lay4 streptococc
13	393	81.0	395	2 Q9LAY2	Q9lay2 streptococc
14	393	81.0	408	2 Q9LAY0	Q9lay0 streptococc
15	392	80.8	426	2 Q9LAY5	Q9lay5 streptococc
16	379.5	78.2	869	2 Q9KK27	Q9kk27 streptococc
17	363.5	74.9	225	2 Q9L591	Q9l591 streptococc
18	357.5	73.7	222	2 Q9L577	Q9l577 streptococc
19	357.5	73.7	262	2 Q9L576	Q9l576 streptococc
20	357.5	73.7	415	2 Q9LAY7	Q9lay7 streptococc
21	354.5	73.1	246	2 Q9L578	Q9l578 streptococc
22	350.5	72.3	416	2 Q9LAY8	Q9lay8 streptococc
23	349.5	72.1	255	2 Q9L581	Q9l581 streptococc
24	349.5	72.1	255	2 Q9L586	Q9l586 streptococc
25	348.5	71.9	406	2 Q9LAZ0	Q9laz0 streptococc
26	343.5	70.8	393	2 Q9LAZ3	Q9laz3 streptococc
27	342.5	70.6	394	2 Q9LAY6	Q9lay6 streptococc
28	342.5	70.6	395	2 Q9LAZ1	Q9laz1 streptococc
29	337.5	69.6	340	2 Q8KQK5	Q8kk5 streptococc
30	334.5	69.0	207	2 Q8GNS9	Q8gns9 streptococc
31	332.5	68.6	237	2 Q9L592	Q9l592 streptococc

ALIGNMENTS

RESULT 1

Q9LAY3 PRELIMINARY; PRT; 417 AA.
AC Q9LAY3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE PcpA (Fragment).
GN Name=pppA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF10197;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.B.;
RT "Diversity of PcpA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071812; AAF27708.1; -
DR HSSP; P00192; 256B. 417
FT NON_TER 417 417
SQ SEQUENCE 417 AA; 46960 MW; 876EAD3276506EEC CRC64;

Query Match 100.0%; Score 485; DB 2; Length 417;
Best Local Similarity 100.0%; Pred. No. 6,5e-23;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKGFRAPLQSELDKQAKLSKLEELSDKIDELDAETAKLEDQLKAAE 60
DB 213 LKEIDSESDYVKGFRAPLQSELDKQAKLSKLEELSDKIDELDAETAKLEDQLKAAE 272

QY 61 ENNVEDYFKEGKLEKTIKAAKLEKTEADLKAVNEPE 99
DB 273 ENNVEDYFKEGKLEKTIKAAKLEKTEADLKAVNEPE 311

RESULT 2

Q54972 PRELIMINARY; PRT; 619 AA.
AC Q54972;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Pneumococcal surface protein A precursor.
GN Name=pppA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]

```

RP SEQUENCE FROM N.A.
RX MEDLINE=92105030; PubMed=1729249;
RA Yother J., Briles D.E.;
RT "Structural properties and evolutionary relationships of PspA, a
RT surface protein of Streptococcus pneumoniae, as revealed by sequence
RT analysis.";
RL J. Bacteriol. 174:601-609(1992).
RN [2]
RN SEQUENCE FROM N.A.
RA Yother J., Briles D.E.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M74122; AAA27018.1; -.
DR PIR; A41971; A41971.
DR PIR; A97887; A97887.
DR HSP; P06653; IHCX.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam; PF01473; CW_binding_1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_3.
KW Signal.
FT SIGNAL 1 31 Potential.
FT CHAIN 32 619 pneumococcal surface protein A.
SQ SEQUENCE 619 AA; 68605 MW; 5AA8BDB40C2841CA CRC64;

Query Match 96.7%; Score 469; DB 2; Length 619;
Best Local Similarity 96.0%; Pred. No. 9.3e-22;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSKDIDELDAETAKLEDQKAAE 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLELSKDIDELDAETAKLEDQKAAE 282
QY 61 ENNVEDYFKGLEKTIAAKAELEKTEADLKKAVNEPE 99
Db 283 ENNVEDYFKGLEKTIAAKAELEKTEADLKKAVNEPE 321

RESULT 3
Q8DR10 PRELIMINARY; PRT; 619 AA.
AC Q8DR10;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein pspA.
GN Name=pspA; OrderedLocustNames=spr0121;
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=171101;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=21429245; PubMed=11544234;
RX DOI=10.1128/JB.183.19.5709-5717.2001;
RA Hoskins J., Albom W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmour R., Glaes J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAhren S.M., McHenry M., Mcleaster K., Mundy C.W., Niclas T.I.,
RA Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rockey P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostock P.R. Jr., Skatrud P.L.,
RA "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RT J. Bacteriol. 183:5709-5717(2001).
RL EMBL; AE008396; AAK98925.1; -.
DR PIR; A41971; A41971.
DR PIR; A97887; A97887.
DR HSP; P06653; IHCX.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam; PF01473; CW_binding_1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_3.

```

```

KW Complete proteome.
SQ SEQUENCE 619 AA; 68605 MW; 5AA8BDB40C2841CA CRC64;

Query Match 96.7%; Score 469; DB 2; Length 619;
Best Local Similarity 96.0%; Pred. No. 9.3e-22;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSKDIDELDAETAKLEDQKAAE 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLELSKDIDELDAETAKLEDQKAAE 282
QY 61 ENNVEDYFKGLEKTIAAKAELEKTEADLKKAVNEPE 99
Db 283 ENNVEDYFKGLEKTIAAKAELEKTEADLKKAVNEPE 321

RESULT 4
Q9LAY1 PRELIMINARY; PRT; 415 AA.
AC Q9LAY1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=1313;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=WU2;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071814; AAF27710.1; -.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
DR NON_TER 415 415
SQ SEQUENCE 415 AA; 46075 MW; 213C1AF7FF21642F CRC64;

Query Match 92.2%; Score 447; DB 2; Length 415;
Best Local Similarity 91.9%; Pred. No. 1.6e-20;
Matches 91; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSKDIDELDAETAKLEDQKAAE 60
Db 221 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLELSKDIDELDAETAKLEDQKAVE 280
QY 61 ENNVEDYFKGLEKTIAAKAELEKTEADLKKAVNEPE 99
Db 281 ENNVEDYFKGLEKTIAAKAELEKTEADLKKAVNEPE 319

RESULT 5
Q9RQT4 PRELIMINARY; PRT; 739 AA.
AC Q9RQT4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein pspC (Fragment).
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=1313;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=E134;

```

```

RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, PspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; AF068647; AAF13457.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; GPos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW_binding_1; 1.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
KW Hypothetical protein.
FT NON_TER 739
SQ SEQUENCE 739 AA; 83960 MW; 7EE2F2F676ABF989 CRC64;

Query Match 85.4%; Score 414; DB 2; Length 739;
Best Local Similarity 86.9%; Pred. No. 3e-18;
Matches 86; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEDQLKAAE 60
DB 537 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEDQLKAAE 596
QY 61 ENNVVDYFKEGLEKTTAAKAELEKTEADLKKAVNEPE 99
DB 597 GNNVVEAYFKEGLEKTTAAKAELEKAEADLKKAVDEPE 635

RESULT 6
Q9RQT1 PRELIMINARY; PRT; 820 AA.
AC Q9RQT1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein pspC (Fragment).
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9163;
RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, PspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; AF068650; AAF13460.1; -.
DR HSSP; P04268; 1IC2.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; GPos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW_binding_1; 1.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
KW Hypothetical protein.
FT NON_TER 820
SQ SEQUENCE 820 AA; 91752 MW; 33C095849ABB0942 CRC64;

Query Match 85.4%; Score 414; DB 2; Length 820;
Best Local Similarity 86.9%; Pred. No. 3.3e-18;
Matches 86; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

```

```

QY 1 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEDQLKAAE 60
DB 530 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEDQLKAAE 589
QY 61 ENNVVDYFKEGLEKTTAAKAELEKTEADLKKAVNEPE 99
DB 590 GNNVVEAYFKEGLEKTTAAKAELEKAEADLKKAVDEPE 628

RESULT 7
Q9KK19 PRELIMINARY; PRT; 929 AA.
AC Q9KK19;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein PspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=srf10;
RX MEDLINE=21888621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae.";
RL Gene 284:63-71(2002).
DR EMBL; AF154037; AAF73809.1; -.
DR HSSP; P06653; 1H8G.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; GPos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW_binding_1; 1.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
SQ SEQUENCE 929 AA; 105003 MW; 2DC8293302FAFA64 CRC64;

Query Match 85.4%; Score 414; DB 2; Length 929;
Best Local Similarity 86.9%; Pred. No. 3.7e-18;
Matches 86; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEDQLKAAE 60
DB 530 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEDQLKAAE 589
QY 61 ENNVVDYFKEGLEKTTAAKAELEKTEADLKKAVNEPE 99
DB 590 GNNVVEAYFKEGLEKTTAAKAELEKAEADLKKAVDEPE 628

RESULT 8
Q9ZAY5 PRELIMINARY; PRT; 929 AA.
AC Q9ZAY5;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein C.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF6796;

```

```

RX MEDLINE=20038319; PubMed=10569772;
RA Brooke-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, pspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; U72655; AAD00184.1; -.
DR HSP; P06653; 1HCX.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_YSRK.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; YSRK signal; 1.
DR TIGRFAMs; TIGR01168; YSRK_signal; 1.
SQ SEQUENCE 929 AA; 104991 MW; 2DC8293302FFB081 CRC64;

Query Match 85.4%; Score 414; DB 2; Length 929;
Best Local Similarity 86.9%; Pred. No. 3.7e-18;
Matches 86; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 LKEIDESDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLELDQKAAE 60
Db 530 LKEIDESDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLELDQKAAE 589

Qy 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKKAVNEPE 99
Db 590 GNNVVEAYFKEGLEKTTAAKKAELKTEADLKKAVDEPE 628

RESULT 9
Q8KQK4
ID Q8KQK4 PRELIMINARY; PRT; 99 AA.
AC Q8KQK4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Pneumococcal surface protein A (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=371/00;
RX MEDLINE=22170754; PubMed=12183557;
DOI=10.1128/IAI.70.9.5086-5090.2002;
RA Miyaji E.N., Ferreira D.M., Lopes A.P.Y., Brandileone M.C.C.,
RA Dias W.O., Leite L.C.C.;
RT "Analysis of serum cross-reactivity and cross-protection elicited by
RT immunization with DNA vaccines against Streptococcus pneumoniae
RT expressing PspA fragments from different clades.";
RL Infect. Immun. 70:5086-5090(2002).
DR EMBL; AY082388; AAL92493.1; -.
DR NON_TER 1
FT NON_TER 99
FT SEQUENCE 99 AA; 11105 MW; 7A13308CA174A3A7 CRC64;

Query Match 83.5%; Score 405; DB 2; Length 99;
Best Local Similarity 85.9%; Pred. No. 1.8e-18;
Matches 85; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

Qy 1 LKEIDESDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLELDQKAAE 60
Db 1 LKEIDESDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAEVLQKDAE 60

Qy 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKKAVNEPE 99
Db 61 GNNVVEAYFKEGLEKTTAAKKAELKTEADLKKAVNEPE 99

```

```

RESULT 10
Q9L575
ID Q9L575 PRELIMINARY; PRT; 249 AA.
AC Q9L575;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=195;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Packiam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=195;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255552; AAF68105.1; -.
FT NON_TER 1
FT NON_TER 249
FT SEQUENCE 249 AA; 26986 MW; 7916D5014E387BD8 CRC64;

Query Match 83.5%; Score 405; DB 2; Length 249;
Best Local Similarity 84.8%; Pred. No. 4.2e-18;
Matches 84; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy 1 LKEIDESDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLELDQKAAE 60
Db 74 LKEIDESDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAEVLQKDAE 133

Qy 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKKAVNEPE 99
Db 134 GNNVVEAYFKEGLEKTTAAKKAELKTEADLKKAVDEPE 172

RESULT 11
Q8GNS8
ID Q8GNS8 PRELIMINARY; PRT; 224 AA.
AC Q8GNS8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PN124;
RX MEDLINE=22241996; PubMed=12354862;
RA Dicuonzo G., Gherardi G., Gertz R.E., D'Ambrosio F., Goglio A.,
RA Lorino G., Recchia S., Pantosti A., Beall B.;
RT "genotypes of invasive pneumococcal isolates recently recovered from
RT Italian patients.";
RL J. Clin. Microbiol. 40:3660-3665(2002).
DR EMBL; AF490267; AAN37735.1; -.
DR HSP; P00192; 1APC.
DR InterPro; IPR009082; His_kin_homodim.
FT NON_TER 1
FT NON_TER 224

```

```
SQ SEQUENCE 224 AA; 23418 MW; 48674E27AFB66A95 CRC64;

Query Match 82.1%; Score 398; DB 2; Length 224;
Best Local Similarity 83.8%; Pred. No. 1e-17;
Matches 83; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYKGFAPLOSLEDAKQAKLSKLELSKDIDELDAETAKLEDOLKAAE 60
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 17 LKIDNESDSYKGFAPLOSLEDAKQAKLSKLELSKDIDELDAETAKLEDOLKAAE 76
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 ENNVEDYFKEGLEKTTAAKAELEKTEADLKAVNEPE 99
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 77 GNNVVEYFKEGLEKTTAEKAELEKAEADLKAVDEPE 115

RESULT 12
Q9LAY4
ID Q9LAY4 PRELIMINARY; PRT; 437 AA.
AC Q9LAY4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RC STRAIN=EL34;
RX MEDLINE=20448953; PubMed=10992499;
RA DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071811; AAF27707.1; -.
DR InterPro; IPR002479; CW binding.
DR Pfam; PF01473; CW binding_1; 1.
FT NON TER 437
SQ SEQUENCE 437 AA; 48071 MW; F6EFD2CD13E08CD8 CRC64;

Query Match 82.1%; Score 398; DB 2; Length 437;
Best Local Similarity 84.8%; Pred. No. 1.9e-17;
Matches 84; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYKGFAPLOSLEDAKQAKLSKLELSKDIDELDAETAKLEDOLKAAE 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 235 LKEIDSDSDYKGFAPLOSLEDAKQAKLSKLELSKDIDELDAETAKHVVLQKAAE 294
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 ENNVEDYFKEGLEKTTAAKAELEKTEADLKAVNEPE 99
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 295 GNNVVEYFKEGLEKTTAEKAELEKAEADLKAVDEPE 333

RESULT 13
Q9LAY2
ID Q9LAY2 PRELIMINARY; PRT; 395 AA.
AC Q9LAY2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RC STRAIN=EF6796;
RX MEDLINE=20448953; PubMed=10992499;

RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071815; AAF27711.1; -.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin_alc_DH_like.
FT NON TER 408
SQ SEQUENCE 408 AA; 44254 MW; 4F64D874217297EF CRC64;

Query Match 81.0%; Score 393; DB 2; Length 408;
Best Local Similarity 82.8%; Pred. No. 3.7e-17;
Matches 82; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYKGFAPLOSLEDAKQAKLSKLELSKDIDELDAETAKLEDOLKAAE 60
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 228 LBEINESDSYKGFAPLOSLEDAKQAKLSKLELSKDIDELDAETAKLEDOLKAAE 287
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 ENNVEDYFKEGLEKTTAAKAELEKTEADLKAVNEPE 99
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 288 GNNVVEYFKEGLEKTTAEKAELEKAEADLKAVDEPE 326

RESULT 15
Q9LAY5
ID Q9LAY5 PRELIMINARY; PRT; 426 AA.
AC Q9LAY5;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
```

```
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBL5;
RX MEDLINE=20448953; PubMed=10992499;
RA DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071810; AAF27706.1; -.
DR HSSP; P00192; IMGT.
DR InterPro; IPR011047; Quin_abc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 426
SQ SEQUENCE 426 AA; 46534 MW; 81AA11348CBE6634 CRC64;

Query Match      80.8%; Score 392; DB 2; Length 426;
Best Local Similarity 82.8%; Pred. No. 4.4e-17;
Matches 82; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Qy 1 LKEIDSDSEDYVKEGFRAPLQSELDKQAKLSKLELSDKIDELDAEIAKLEDLKAAE 60
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 215 LKDINESDSEDYVKEGLRAPLQSELDTKAKLLKLELSGKIEELDAEIAELEVLQKDAE 274
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 61 ENNVEDYFKEGLEKTTAAKKAELEKTEADLKKAVNEPE 99
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 275 GNNVVEAYFKEGLEKTTAEKKAELEKAEADLUKKAVDPE 313
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Search completed: June 21, 2005, 10:22:10
Job time : 63.3194 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:50:39 ; Search time 74.5918 Seconds
(without alignments)

518.502 Million cell updates/sec

Title: US-10-674-755-12

Perfect score: 489

Sequence: 1 LKIDSESDYAKGFRAP.....KKALEKTEADLKAVNEPE 100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	469.5	96.0	198	7	ABW02615 Rxlc pneu
2	469.5	96.0	315	2	AAY04375 Streptoco
3	469.5	96.0	619	2	AAR63437 Pneumococ
4	469.5	96.0	619	2	AAR87598 Pneumococ
5	469.5	96.0	619	2	AAR86911 Pneumococ
6	469.5	96.0	619	2	AAY41838 Streptoco
7	469.5	96.0	619	5	AAEL8782 S. pneumo
8	469.5	96.0	619	6	ABU45778 Protein e
9	469.5	96.0	619	8	ADOS2126 Streptoco
10	469.5	96.0	648	2	AAW70336 Pneumococ
11	469.5	96.0	648	2	AAW62274 Streptoco
12	469.5	96.0	648	2	AAY41837 Streptoco
13	469.5	96.0	648	2	AAW87879 A pneumoc
14	469.5	96.0	653	2	AAW92456 S. pneumo
15	469.5	96.0	684	2	AAAR73912 Streptoco
16	469.5	96.0	8991	6	ABU08487 S. pneumo
17	466.5	95.4	198	2	AAW14581 Streptoco
18	462.5	94.6	204	2	AAW14571 Streptoco
19	462.5	94.6	204	7	ABW02605 Efi019c p
20	449.5	91.9	653	2	AAAR27150 PspA frag
21	447	91.4	289	2	AAW62276 Streptoco
22	447	91.4	289	2	AAY41840 Streptoco
23	447	91.4	289	2	AAW87910 Protein s
24	447	91.4	289	2	AAW92458 S. pneumo
25	442.5	90.5	195	2	AAW14591 Streptoco

26	442.5	90.5	195	7	ABW02625	Abw02625 Wu2c pneu
27	426.5	87.2	1231	6	ABU08490	Abu08490 Fragment
28	423.5	86.6	623	6	ABU08494	Abu08494 Fragment
29	402.5	82.3	170	7	ABW02614	Abw02614 Rct135c p
30	402.5	82.3	181	7	ABW02596	Abw02596 0922134c
31	402.5	82.3	865	6	ABU08489	Abu08489 S. pneumo
32	402.5	82.3	929	2	AAW14593	AAW14593 Streptoco
33	402.5	82.3	929	2	AAY43384	Aay43384 S. pneumo
34	399.5	81.7	188	7	AAW14580	AAW14580 Streptoco
35	399.5	81.7	188	7	ABW02613	Abw02613 Rct129c p
36	392.5	80.3	204	2	AAW14578	AAW14578 Streptoco
37	392.5	80.3	204	7	ABW02612	Abw02612 Rct123c p
38	390.5	79.9	588	6	ABU08491	Abu08491 Coiled co
39	390.5	79.9	589	2	AAY43392	Aay43392 PspC alph
40	388	79.3	180	2	AAW14562	AAW14562 Streptoco
41	385	78.7	187	2	AAW14579	AAW14579 Streptoco
42	375.5	76.8	206	2	AAW14574	AAW14574 Streptoco
43	375.5	76.8	206	7	ABW02608	Abw02608 Db15c pne
44	351	71.8	550	8	ADK48356	Adk48356 Streptoco
45	351	71.8	550	8	ADR95223	Adr95223 Novel S.

ALIGNMENTS

RESULT 1

ABW02615

ID ABW02615 standard; protein; 198 AA.

AC ABW02615;

DT 12-FEB-2004 (first entry)

DE Rxlc pneumococcal surface protein A (PspA) central region.

KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;

KW immunological; gene therapy; immunostimulant.

OS Unidentified.

PN US6592876-B1.

PD 15-JUL-2003.

PF 15-SEP-1995; 95US-00529055.

PR 20-APR-1993; 93US-00048896.

PR 06-JUN-1995; 95US-00465746.

XX (UABR-) UAB RES FOUND.

XX Briles DB, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;

XX WPI; 2003-862841/80.

XX Immunological composition for obtaining expression products used for

XX detecting the presence of Streptococcus pneumoniae or its strain,

XX comprises at least two different full length isolated gene encoding

XX pneumococcal surface protein A.

XX Example 6; SEQ ID NO 61; 121pp; English.

XX The present invention relates to an immunological composition comprising

XX at least 2 different full length isolated genes encoding pneumococcal

XX surface protein A (PspAs) from different groups based on restriction

XX fragment polymorphism analysis. The invention is useful for obtaining

XX expression products by recombinant techniques to detect, determine,

XX isolate or diagnose the presence of Streptococcus pneumoniae or its

XX strain. The expression product is useful for preparing antigenic,

XX immunological or vaccine compositions, for eliciting antibodies, an

XX immunological response (other than or additional to antibodies) or a

XX protective response (including antibody or other immunological response

XX by administering compositions to a host). The invention is also useful as

CC vaccines and in gene therapy. The present sequence is RxlC pneumococcal
 CC surface protein A (PspA) central region. This sequence is used in the
 CC exemplification of the invention

XX SQ Sequence 198 AA;

Query Match 96.0%; Score 469.5; DB 7; Length 198;
 Best Local Similarity 98.0%; Pred. No. 4.7e-34;
 Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLELSKIDELDAEIAKLEDLQKAA 60
 |||||
 Db 1 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLELSKIDELDAEIAKLEDLQKAA 59
 |||||

Qy 61 EENNVEDYFKGLEKTIIAAKKAELEKTEADLKAVNEPE 100
 |||||

Db 60 EENNVEDYFKGLEKTIIAAKKAELEKTEADLKAVNEPE 99
 |||||

RESULT 2

ID AAY04375 standard; protein; 315 AA.

XX AAY04375;

DT 23-JUN-1999 (first entry)

DE Streptococcus pneumoniae PspA protein sequence.

KW Streptococcus pneumoniae; pspA; pneumococcal; surface protein; vaccine;
 KW immunological; infection.

OS Streptococcus pneumoniae.

OS Synthetic.

XX WO9914333-A2.

PN 25-MAR-1999.

PD 18-SEP-1998; 98WO-US019740.

PF 18-SEP-1997; 97US-00932982.

PR (INNR) PASTEUR MERIEUX CONNAUGHT.

PA Becker R, Gray M, Pyle D;

PI WPI; 1999-229537/19.

DR N-PSDB; AAX33124.

XX DNA encoding PspA molecule with modified internal translational
 PT initiation sites.

XX Disclosure; Page; 36pp; English.

XX The present sequence represents a pneumococcal surface protein A (PspA)
 CC molecule where internal naturally occurring translational initiation
 CC sites have been modified or eliminated so that expression of the DNA
 CC sequence results in a single form of PspA. The PspA nucleotide sequence
 CC can be used to transform a unicellular host to produce the PspA protein.
 CC The PspA protein can be used in an immunological composition for treating
 CC or preventing S. pneumoniae infection especially in a child. Antibodies
 CC to the PspA protein can also be used to treat S. pneumoniae infection.
 CC The immunogenic peptides are designed to confer broad protection against
 CC diverse pneumococcal strains. N.B. The present sequence is not given in
 CC the specification but is encoded by the sequence given in AAX33124

XX SQ Sequence 315 AA;

Query Match 96.0%; Score 469.5; DB 2; Length 315;
 Best Local Similarity 98.0%; Pred. No. 8.2e-34;
 Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLELSKIDELDAEIAKLEDLQKAA 60
 |||||
 Db 193 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLELSKIDELDAEIAKLEDLQKAA 251
 |||||

Qy 61 EENNVEDYFKGLEKTIIAAKKAELEKTEADLKAVNEPE 100
 |||||

Db 252 EENNVEDYFKGLEKTIIAAKKAELEKTEADLKAVNEPE 291
 |||||

RESULT 3

AAR63437

ID AAR63437 standard; protein; 619 AA.

XX AAR63437;

AC 09-SEP-2004 (revised)

DT 16-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 19-JUL-1995 (first entry)

XX Pneumococcal surface protein A from S.pneumoniae Rxl.

XX Pneumococcal surface protein A; pspA; Streptococcus; PCR; pneumococcal;
 KW primer; protection-eliciting epitope; epitope; vaccine; amplify.

OS Streptococcus pneumoniae.

OS Unidentified.

XX Key Location/Qualifiers

FT Protein 192..260

FT /note= "protein fragment of Claim 1"

XX EP622081-A2.

PN 02-NOV-1994.

PD 19-APR-1994; 94EP-00302767.

XX 20-APR-1993; 93US-00048896.

XX (UABR-) UAB RES FOUND.

XX Briles DE, Yother JL, McDaniel LS;

XX WPI; 1994-359522/45.

XX N-PSDB; AAO78131.

XX regions of Pneumococcal surface protein A - derived from the Rxl PspA
 PT strain, for the preparation of cross-reactive vaccines for the prevention
 PT of pneumococcal infections.

PS Disclosure; Page 13-16; 26pp; English.

XX The amino acid sequence of the novel Pneumococcal surface protein A
 CC (PspA) from Streptococcus pneumoniae strain Rxl. The gene was PCR
 CC amplified from S.pneumoniae genomic DNA using the primers AAO78132-5. The
 CC gene was used to derive truncated peptide fragments containing protection
 CC -eliciting epitopes for use in vaccines against pneumococcal diseases.
 CC The epitopic fragments are derived from amino acids 192-260 and C-
 CC optionally contain a further 25 a.a. residues at both the N- and C-
 CC terminal regions of the peptide. The epitopic peptide fragments may be
 CC derived from different strains of S.pneumoniae and are homologous to the
 CC Rxl strain epitope. (Updated on 25-MAR-2003 to correct PN field.)
 CC (Updated on 16-OCT-2003 to standardise OS field)

CC Revised record issued on 09-SEP-2004 : Correction to feature table key

XX SQ Sequence 619 AA;

Query Match 96.0%; Score 469.5; DB 2; Length 619;
 Best Local Similarity 98.0%; Pred. No. 1.9e-33;
 Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

```
QY 1 LKEIDSESEDYAKGFRAPLOSKLDKAKOAKLSKLELSKIDELDAETAKLEDOLKAA 60
Db 223 LKEIDSESEDYAKGFRAPLOSKLDKAKKAKLSKLELSKIDELDAETAKLEDOL-KAA 281
QY 61 EENNVEDYFKGEGLEKTIAAKKAELKTEADLKAVNEPE 100
Db 282 EENNVEDYFKGEGLEKTIAAKKAELKTEADLKAVNEPE 321

RESULT 4
AAR87598
ID AAR87598 standard; protein; 619 AA.
XX
AC AAR87598;
XX
XX 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 04-JUL-1996 (first entry)
XX
DE Pneumococcal surface protein (PspA).
XX
XX PspA; pneumococcal surface protein; truncated; immunoprotective;
KW soluble fragment; insertion-duplication mutagenesis.
XX
XX Streptococcus pneumoniae; strain Rxl.
XX
FH Key Location/Qualifiers
FT Peptide 1..31
FT Protein 32..619 /label= signal_peptide
FT Region 32..319 /label= mature_protein
FT /label= alpha-helical coiled-coil region
FT /note= "contains a seven-residue periodicity"
FT Region 320..401
FT Region 402..421 /note= "proline-rich region"
FT Region 422..441 /note= "repeat region"
FT Region 442..461 /note= "repeat region"
FT Region 462..481 /note= "repeat region"
FT Region 482..501 /note= "repeat region"
FT Region 502..521 /note= "repeat region"
FT Region 522..541 /note= "repeat region"
FT Region 542..561 /note= "repeat region"
FT Region 562..581 /note= "repeat region"
FT Region 582..619 /note= "repeat region"
FT /note= "hydrophobic region starts in last repeat region
FT is potential membrane-spanning region"
XX
XX US5476929-A.
XX
XX 19-DEC-1995.
XX
XX 03-JUN-1993; 93US-00072070.
XX
XX 15-FEB-1991; 91US-00656773.
XX 12-FEB-1992; 92US-00835698.
XX
XX (UABR-) UAB RES FOUND.
XX
XX McDaniel LS, Yother JL, Briles DE;
XX
XX WPI; 1996-049021/05.
DR N-PSDB; AAT08979.
```

```
XX New pneumococcal surface protein A fragments - comprise proline-rich
PT region and/or repeat region, used for detection of Streptococcus
PT pneumoniae.
XX
XX Claim 1; Col 15-20; 23pp; English.
XX
XX The present sequence is that of PspA (pneumococcal surface protein A)
CC encoded by AAT08979. Through the technique of insertion-duplication
CC mutagenesis of the pspA gene of the strain Rxl of Streptococcus
CC pneumoniae with plasmids contg. cloned fragments of the pspA structural
CC gene, it has been possible to produce soluble fragments of PspA that are
CC secreted by pneumococci. The method can be used to provide an
CC immunoprotective truncated PspA protein. Primers and probes based on the
CC present sequence are claimed, and are useful for the detection of (at
CC least 32) S. pneumoniae strains. (Updated on 25-MAR-2003 to correct PF
CC field.) (Updated on 16-OCT-2003 to standardise OS field)
XX
XX Sequence 619 AA;
XX
XX Query Match 96.0%; Score 469.5; DB 2; Length 619;
XX Best Local Similarity 98.0%; Pred. NO. 1.9e-33;
XX Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 1 LKEIDSESEDYAKGFRAPLOSKLDKAKOAKLSKLELSKIDELDAETAKLEDOLKAA 60
Db 223 LKEIDSESEDYAKGFRAPLOSKLDKAKKAKLSKLELSKIDELDAETAKLEDOL-KAA 281
QY 61 EENNVEDYFKGEGLEKTIAAKKAELKTEADLKAVNEPE 100
Db 282 EENNVEDYFKGEGLEKTIAAKKAELKTEADLKAVNEPE 321

RESULT 5
AAR86911
ID AAR86911 standard; protein; 619 AA.
XX
AC AAR86911;
XX
XX 16-OCT-2003 (revised)
DT 11-MAY-1996 (first entry)
DE Pneumococcal surface protein A.
XX
XX Pneumococcal surface protein A; PspA; cross-protection; vaccine;
KW Streptococcus pneumoniae; probe; primer; polymerase chain reaction;
KW otitis media; meningitis; bacteraemia; pneumonia; epitope.
XX
XX Streptococcus pneumoniae; strain Rxl.
XX
XX Key Location/Qualifiers
FT Peptide 1..31
FT /label= sig_peptide
FT Region 32..288
FT /note= "N-terminal region is highly charged and includes
FT an alpha-helix structure"
FT 289..619
FT Region /note= "C-terminal region includes a proline-rich region
FT and a repeat region"
XX
XX AU9520112-A.
XX
XX 30-NOV-1995.
XX
XX 18-MAY-1995; 95AU-00020112.
XX
XX 20-MAY-1994; 94US-00246636.
XX 07-OCT-1994; 94US-00319795.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, Yother JL, McDaniel LS;
XX
XX
```

```

DR WPI; 1996-030801/04.
DR N-PSDB; AAT07178.
XX Pneumococcal DNA primers and probes - amplify and detect cross-protective
PT epitope(s) from Streptococcus pneumoniae surface protein A.
XX Disclosure; Page 41-43; 61pp; English.
XX Surface protein A, PspA (AAR86911), of Streptococcus pneumoniae Rx1 is
CC the product of the pspA gene (AAT07178). PspA includes regions
CC comprising e.g. amino acids 182-588, 293-588 and 192-299, that elicit
CC cross-protection against challenge by multiple wild-type strains of S.
CC pneumoniae. These cross-reactive epitopes can be prepd. by expression of
CC DNA obt'd. by PCR amplification (see AAT07179-96), for use in vaccine
CC compns. (Updated on 16-OCT-2003 to standardise OS field)
XX Sequence 619 AA;

Query Match          96.0%; Score 469.5; DB 2; Length 619;
Best Local Similarity 98.0%; Pred. No. 1.9e-33;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Oy 1 LKEIDSESEDYAKGFRAPLQSKLDKAKQAKLSKLELSKIDELDAEIAKLEDLQKAA 60
Db 223 LKEIDSESEDYAKGFRAPLQSKLDKAKQAKLSKLELSKIDELDAEIAKLEDLQKAA 281

Oy 61 EENNVEDYFKGLEKTIAAKAELEKTEADLKAVNEPE 100
Db 282 EENNVEDYFKGLEKTIAAKAELEKTEADLKAVNEPE 321

RESULT 6
AAAY41838
ID AAAY41838 standard; protein; 619 AA.
XX AC AAAY41838;
XX DT 08-DEC-1999 (first entry)
XX DE Streptococcus pneumoniae Rx1 PspA protein sequence.
XX KW Streptococcus pneumoniae Rx1; PspA; immunoprotective; vaccine; diagnosis;
XX infection; pneumococcal surface protein A.
XX OS Streptococcus pneumoniae.
XX PN US965400-A.
XX PD 12-OCT-1999.
XX PF 23-MAY-1994; 94US-00247491.
XX PR 15-FEB-1991; 91US-00656773.
XX PR 12-FEB-1992; 92US-00835698.
XX PA (UABR-) UAB RES FOUND.
XX PI Yother JL, Briles DE;
XX WPI; 1999-579913/49.
DR N-PSDB; AAZ25063.
XX DNA encoding a truncated pneumococcal surface protein A used in the
XX development of pneumococcal infections.
XX Claim 1; Fig 3; 27pp; English.
XX The present sequence represents Streptococcus pneumoniae Rx1
CC immunoprotective Pneumococcal surface protein A (PspA). The present
CC invention also describes a method of forming the immunoprotective
CC truncated PspA, comprising incorporating a vector comprising the isolated
CC DNA molecule encoding PspA (I), into a bacterium via transformation. (I)
CC is used to design primers which are capable of detecting a large number

```

```

CC of S. pneumoniae strains, which in turn can be used to diagnose
CC pneumococcal infection in mammals (e.g. humans), independent of the
CC strain which has caused it. The PspA protein is used to develop a vaccine
CC against pneumococcal infection comprising, as an immunologically-active
CC component, a live attenuated or killed bacteria containing a gene coding
CC for the truncated form of PspA
XX Sequence 619 AA;

Query Match          96.0%; Score 469.5; DB 2; Length 619;
Best Local Similarity 98.0%; Pred. No. 1.9e-33;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Oy 1 LKEIDSESEDYAKGFRAPLQSKLDKAKQAKLSKLELSKIDELDAEIAKLEDLQKAA 60
Db 223 LKEIDSESEDYAKGFRAPLQSKLDKAKQAKLSKLELSKIDELDAEIAKLEDLQKAA 281

Oy 61 EENNVEDYFKGLEKTIAAKAELEKTEADLKAVNEPE 100
Db 282 EENNVEDYFKGLEKTIAAKAELEKTEADLKAVNEPE 321

RESULT 7
AAE18782
ID AAE18782 standard; protein; 619 AA.
XX AC AAE18782;
XX DT 17-MAY-2002 (first entry)
XX DE Streptococcus pneumoniae.
XX KW Streptococcus pneumoniae Rx1 strain pneumococcal surface protein A (PspA).
XX cell-mediated immunity; microbial infection; cross-protection; therapy;
XX antimicrobial; vaccine; pneumococcal surface protein A; PspA.
XX OS Streptococcus pneumoniae.
XX PH Key Location/Qualifiers
FT Domain 1..314 /label= Helical_domain
FT Region 1..303 /note= "N-terminal region"
FT Region 38..44 /note= "Immunogenic region 3"
FT Region 40..46 /note= "Immunogenic region 5"
FT Region 75..80 /note= "Immunogenic region 29"
FT Region 82..87 /note= "Immunogenic region 52"
FT Region 96..101 /note= "Immunogenic region 66"
FT Region 114..119 /note= "Immunogenic region 73"
FT Region 130..135 /note= "Immunogenic region 78"
FT Region 137..142 /note= "Immunogenic region 89"
FT Region 140..145 /note= "Immunogenic region 91"
FT Region 152..156 /note= "Immunogenic region 95"
FT Domain 153..170 /label= Coiled_coil_motif
FT Region 161..164 /note= "Immunogenic region 101"
FT Region 166..170 /note= "Immunogenic region 116"
FT Region 173..177 /note= "Immunogenic region 122"
FT Region 176..180 /note= "Immunogenic region 123"

```


CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX
 SQ Sequence 619 AA;

Query Match 96.0%; Score 469.5; DB 6; Length 619;
 Best Local Similarity 98.0%; Pred. No. 1.9e-33;
 Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLELSDKIDELDAEIAKLEDQKAA 60
 |||||
 DB 223 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLELSDKIDELDAEIAKLEDQK-KAA 281
 |||||

QY 61 EENNVEDYFKGLEKTIAAKKAELKTEADLKKAVNEPE 100
 |||||
 DB 282 EENNVEDYFKGLEKTIAAKKAELKTEADLKKAVNEPE 321
 |||||

RESULT 9
 ADO52126
 ID ADO52126 standard; protein; 619 AA.

XX
 AC ADO52126;
 XX
 DT 12-AUG-2004 (first entry)

XX Streptococcus pneumoniae Rx1 PspA protein.
 DE
 XX
 KW Immunogenic composition; vaccine; Th2-type immune response;
 KW pneumococcal surface protein A; PspA.
 XX

XX Streptococcus pneumoniae.
 FH Key Location/Qualifiers
 FT Peptide 1..31 /label=Signal_peptide
 FT Protein 32..619 /note="S. pneumoniae Rx1 mature PspA protein"
 FT
 FT
 XX US2004101531-A1.
 PN
 XX
 XX 27-MAY-2004.
 PD
 XX
 XX 15-APR-2003; 2003US-00414532.
 PF
 XX
 XX 16-APR-2002; 2002US-0372710P.
 PR
 XX (CURT/) CURTISS R.
 PA (KANG/) KANG H Y.
 XX
 XX Curtiss R, Kang HY;
 PI
 XX WPI: 2004-399655/37.
 DR N-PSDB; ADO52125.
 DR
 XX

XX New vaccine comprising a live attenuated strain of pathogenic gram-
 PT negative bacteria, useful in eliciting a Th2-type immune response in a
 PT vertebrate against pathogens, e.g., helminths, fungi, viruses, protozoans
 PT or bacteria.
 PT

XX Example 8; SEQ ID NO 72; 94pp; English.
 PS
 XX The invention relates to immunogenic compositions and vaccines comprising
 CC a live attenuated strain of pathogenic gram negative bacteria that
 CC secretes an antigen. The vaccine is useful in eliciting a Th2-type immune
 CC response in a vertebrate against pathogens, e.g., helminths, fungi,
 CC viruses, protozoans or bacteria. The present sequence is Streptococcus
 CC pneumoniae Rx1 pneumococcal surface protein A (PspA). This sequence is
 CC used in the exemplification of the invention.

XX
 SQ Sequence 619 AA;

Query Match 96.0%; Score 469.5; DB 8; Length 619;
 Best Local Similarity 98.0%; Pred. No. 1.9e-33;
 Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLELSDKIDELDAEIAKLEDQKAA 60
 |||||
 DB 223 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLELSDKIDELDAEIAKLEDQK-KAA 281
 |||||

QY 61 EENNVEDYFKGLEKTIAAKKAELKTEADLKKAVNEPE 100
 |||||
 DB 282 EENNVEDYFKGLEKTIAAKKAELKTEADLKKAVNEPE 321
 |||||

RESULT 10
 AAW70336
 ID AAW70336 standard; protein; 648 AA.

XX
 AC AAW70336;
 XX
 DT 18-NOV-1998 (first entry)

XX Pneumococcal surface protein A (PspA).
 DE
 XX
 KW Pneumococcal surface protein A gene; PspA; PspA epitope; vaccine;
 KW insertion-duplication mutagenesis; otitis media; meningitis; bacteraemia;
 KW pneumonia.
 XX

XX Streptococcus pneumoniae.
 FH Key Location/Qualifiers
 FT Peptide 1..31 /note="Signal peptide"
 FT Protein 32..648 /note="PspA"
 FT Region 32..319 /note="alpha-helical coil region representing the
 FT truncated PspA of the invention"
 FT Misc-difference 647 /note="Encoded by AGG"
 FT
 XX US5804193-A.
 PN
 XX
 XX 08-SEP-1998.
 PD
 XX
 XX 17-MAR-1994; 94US-00214222.
 PF
 XX
 XX 15-FEB-1991; 91US-00656773.
 PR
 XX 12-FEB-1992; 92US-00835698.
 PR
 XX (UABR-) UAB RES FOUND.
 PA
 XX Briles DE, Yother JL;
 PI
 XX WPI: 1998-505588/43.
 DR N-PSDB; AAW33264.
 DR
 XX

XX Truncated pneumococcal surface protein - useful in vaccines against
 PT pneumococcal infection.
 PT
 XX
 XX Example 3; Fig 3A-3C; 22pp; English.
 PS

XX The present sequence represents the Streptococcus pneumoniae Rx1
 CC pneumococcal surface protein A (PspA). The invention provides a purified
 CC truncated form of PspA, formed by an insertion-duplication mutagenesis
 CC technique, comprising of the first 288 N-terminal residues of the mature
 CC form of wild-type PspA (AAW70336). The truncated PspA contains
 CC immunoprotective epitopes of PspA. The invention claims for a vaccine
 CC against pneumococcal infection, comprising live-attenuated or killed *S.*
 CC pneumoniae, containing the gene coding for the truncated PspA protein.
 CC The truncated protein, optionally conjugated to a poorly immunogenic or

protein from *Streptococcus pneumoniae*. A recombinant DNA molecule has been developed which encodes a fusion protein comprising a truncated form of PspA and cholera toxin B subunit (CTB), where the DNA molecule comprises a nucleotide sequence encoding the truncated PspA linked by an in-frame genetic fusion to a *ctxB* gene, and where the truncated PspA contains immunoprotective epitopes and up to 90% of the whole PspA protein, except for the cell membrane anchor region. The fusion protein is useful for providing an immunogen to protect neonates and children against *S. pneumoniae*. Most antigenic proteins of this strain are not immunogenic enough to provide protection. The antigenic epitopes of the fusion protein are directed against capsular polysaccharide antigens of *S. pneumoniae*, specifically it contains the protective epitopes of PspA. The protein can also be used in solid-phase immunoadsorbent assays, since it is readily bound to supports coated with monosialoganglioside GM1. The fusion protein is more immunogenic against *S. pneumoniae* than using PspA

CC truncated PspA, comprising incorporating a vector comprising the isolated
 CC DNA molecule encoding PspA (I), into a bacterium via transformation. (I)
 CC is used to design primers which are capable of detecting a large number
 CC of *S. pneumoniae* strains, which in turn can be used to diagnose
 CC pneumococcal infection in mammals (e.g. humans), independent of the
 CC strain which has caused it. The PspA protein is used to develop a vaccine
 CC against pneumococcal infection comprising, as an immunologically-active
 CC component, a live attenuated or killed bacteria containing a gene coding
 CC for the truncated form of PspA
 XX
 SQ Sequence 648 AA;

Query Match 96.0%; Score 469.5; DB 2; Length 648;
 Best Local Similarity 98.0%; Pred. No. 2e-33;
 Matches 98; Conservative 1; Mismatches 1; Gaps 1;
 QY 1 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLELSKDIDELDAEIAKLEDQLKAA 60
 DB 223 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLELSKDIDELDAEIAKLEDQLKAA 281
 QY 61 EENNVEDYFKGLEKTIAAKKAELKTEADLKKAVNEPE 100
 DB 282 EENNVEDYFKGLEKTIAAKKAELKTEADLKKAVNEPE 321

RESULT 13
 AA087879
 ID AA087879 standard; protein; 648 AA.

AC AA087879;
 XX
 XX 19-MAR-1999 (first entry)
 XX
 XX A pneumococcal surface protein (PspA).
 XX Pneumococcal surface protein; PspA; Streptococcus pneumoniae Rx1;
 KW truncated protein; epitope-containing region; vaccine;
 KW pneumococcal infection.
 XX Streptococcus pneumoniae.
 XX US5856170-A.
 XX
 XX 05-JAN-1999.
 XX
 XX 06-JUN-1995; 95US-00467852.
 XX
 XX 15-FEB-1991; 91US-00656773.
 XX 12-FEB-1992; 92US-00835698.
 XX 23-MAY-1994; 94US-00247491.
 XX (UABR-) UAB RES FOUND.
 XX
 XX Yother JL, Briles DE;
 XX WPI; 1999-105118/09.
 XX N-PSDB; AAV84069.
 XX DNA encoding truncated pneumococcal PspA protein - useful for producing
 FT recombinant truncated protein.
 XX
 XX Disclosure; Fig 3A-D; 27pp; English.
 XX
 XX The present sequence represents a pneumococcal surface protein (PspA) of
 CC Streptococcus pneumoniae Rx1. The specification describes truncated forms
 CC of PspA which contain at least the N-terminal alpha-helical protective
 CC epitope-containing region. The truncated PspA proteins can be used in
 CC vaccines against pneumococcal infections
 XX
 SQ Sequence 648 AA;

Query Match 96.0%; Score 469.5; DB 2; Length 648;
 Best Local Similarity 98.0%; Pred. No. 2e-33;

Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
 QY 1 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLELSKDIDELDAEIAKLEDQLKAA 60
 DB 223 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLELSKDIDELDAEIAKLEDQLKAA 281
 QY 61 EENNVEDYFKGLEKTIAAKKAELKTEADLKKAVNEPE 100
 DB 282 EENNVEDYFKGLEKTIAAKKAELKTEADLKKAVNEPE 321
 RESULT 14
 AA092456
 ID AA092456 standard; protein; 653 AA.
 XX
 XX AA092456;
 XX
 XX 21-APR-1999 (first entry)
 XX
 XX S. pneumoniae truncated pspA protein.
 XX PspA; pneumococcal surface protein A; immunoprotection; detection; GMI;
 KW solid phase immunosorbent assay; epitope; cell membrane anchor region;
 KW cholera toxin B subunit; CTB; monosialoganglioside; fusion protein.
 XX Streptococcus pneumoniae.
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 620
 FT Misc-difference /note= "in frame stop codon encoded by TAA"
 FT Misc-difference 623
 FT Misc-difference /note= "In frame stop codon encoded by TAA"
 FT Misc-difference 628
 FT Misc-difference /note= "In frame stop codon encoded by TAA"
 FT Misc-difference 634
 FT Misc-difference /note= "in frame stop codon encoded by TAA"
 FT Misc-difference 636
 FT Misc-difference /note= "In frame stop codon encoded by TGA"
 XX
 XX US5871943-A.
 XX
 XX 16-FEB-1999.
 XX
 XX 06-JUN-1995; 95US-00468718.
 XX
 XX 15-FEB-1991; 91US-00656773.
 XX 12-FEB-1992; 92US-00835698.
 XX 03-JUN-1993; 93US-00072068.
 XX (UABR-) UAB RES FOUND.
 XX
 XX Briles DE, Yother JL;
 XX WPI; 1999-166635/14.
 XX N-PSDB; AAX02012.
 XX Immunosorbent assay for pneumococcal surface protein A antigen or
 PT antibody - for diagnosis of infection by Streptococcus pneumoniae.
 XX
 XX Claim 1; Fig 3A-C; 24pp; English.
 XX
 XX This sequence represents a truncated form of the Streptococcus pneumoniae
 CC PspA protein which is used in a solid phase immunosorbent assay for
 CC detecting a PspA (pneumococcal surface protein A) antibody and antigen.
 CC This truncated protein contains the immunoprotective epitopes of the
 CC complete protein (up to 90% of PspA but excludes the cell-membrane anchor
 CC region) fused to the B subunit of cholera toxin (CTB) which is bound to
 CC monosialoganglioside (GMI) coated on the substrate. The use of a fusion
 CC between truncated PspA and cholera toxin B subunit (CTB) allows the
 CC support to be coated without having to isolate PspA fragments since CTB
 CC binds specifically to the GMI coating the solid support
 XX
 XX Sequence 653 AA;

Query Match 96.0%; Score 469.5; DB 2; Length 653;
 Best Local Similarity 98.0%; Pred. No. 2e-33;
 Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESEDYAKGFRAPLOSKLDKAKOAKLSKLEELSDKIDELDAETAKLEDQLKAA 60
 |||||:|||||
 Db 223 LKEIDSESEDYAKGFRAPLOSKLDKAKOAKLSKLEELSDKIDELDAETAKLEDQLKAA 281
 |||||:|||||

QY 61 ENNNVEDYFKGEGLEKTIAAKKAELKTEADLKKAVNEPE 100
 |||||:|||||
 Db 282 ENNNVEDYFKGEGLEKTIAAKKAELKTEADLKKAVNEPE 321
 |||||:|||||

RESULT 15

AAR73912
 ID AAR73912 standard; protein; 684 AA.

AC AAR73912;

XX 25-MAR-2003 (revised)

DT 05-DEC-1995 (first entry)

XX Streptococcus pneumoniae surface protein A.

XX Streptococcus pneumoniae; surface protein A; vaccine;

KW meningitis related homologous antigenic sequence; MRHAS; RV-1;

KW immunoassay; diagnosis; treatment; prophylactic; bacterial; viral.

XX Streptococcus pneumoniae.

XX Key Location/Qualifiers

PH 1. .5

FT Peptide

FT /label= sig_peptide

XX WO9509232-A2.

XX 06-APR-1995.

XX 28-SEP-1994; 94WO-CA000516.

XX 28-SEP-1993; 93US-00127499.

XX (VALS/) VAN ALSTYNE D.

PA (SHAR/) SHARMA L R.

XX Van Alstyne D, Sharma LR;

PI WPI; 1995-147431/19.

DR New peptide(s) and corresp. antibodies for the treatment of meningitis -

XX the peptide(s) corresp. to homologous antigenic sites on bacterial and

PT viral agents and on chemokine(s), used for detecting and preventing

PT meningitis.

XX Claim 47; Fig 6/10; 98pp; English.

PS AAR73912 is the Streptococcus pneumoniae surface protein A. It contains

XX the meningitis related antigenic sequences (MRHAS) claimed in AAR73990

CC and AAR73902, which are recognised by a monoclonal antibody from the

CC hybridoma Rubella virus (RV)-1. The claimed MRHAS peptides may be used in

CC immunoassays to diagnose the presence of bacterial and/or viral

CC meningitis agents in a sample, or in prophylactic and therapeutic

CC meningitis treatments. The peptides may also be used as vaccines against

CC meningitis. NB: Identified by matching corresponding MRHAS peptides.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 684 AA;

SQ Query Match 96.0%; Score 469.5; DB 2; Length 684;

Best Local Similarity 98.0%; Pred. No. 2.1e-33;

Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESEDYAKGFRAPLOSKLDKAKOAKLSKLEELSDKIDELDAETAKLEDQLKAA 60
 |||||:|||||
 Db 259 LKEIDSESEDYAKGFRAPLOSKLDKAKOAKLSKLEELSDKIDELDAETAKLEDQLKAA 317
 |||||:|||||

QY 61 ENNNVEDYFKGEGLEKTIAAKKAELKTEADLKKAVNEPE 100
 |||||:|||||

Db 318 ENNNVEDYFKGEGLEKTIAAKKAELKTEADLKKAVNEPE 357
 |||||:|||||

Search completed: June 21, 2005, 10:10:13

Job time : 74.5918 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:55:14 ; Search time 18.6735 Seconds
(without alignments)
399.760 Million cell updates/sec

Title: US-10-674-755-12
Perfect score: 489
Sequence: 1 LKEIDSESEDYAKGFRAP.....KKAELKTEADLKKA VNEPE 100

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A COMB pep.*
2: /cgn2_6/prodata/1/iaa/5B COMB pep.*
3: /cgn2_6/prodata/1/iaa/6A COMB pep.*
4: /cgn2_6/prodata/1/iaa/6B COMB pep.*
5: /cgn2_6/prodata/1/iaa/PC TUS COMB pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1 pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

				SUMMARIES	
Result No.	Score	Query Match	Length DB ID	Description	
1	489	100.0	100	4	US-09-147-875A-12
2	469.5	96.0	99	2	US-08-710-749-11
3	469.5	96.0	198	4	US-08-529-055-61
4	469.5	96.0	619	1	US-08-465-746-2
5	469.5	96.0	619	1	US-08-214-164-2
6	469.5	96.0	619	2	US-08-467-852A-3
7	469.5	96.0	619	2	US-08-246-636-2
8	469.5	96.0	619	2	US-08-247-491A-3
9	469.5	96.0	619	2	US-08-319-795-2
10	469.5	96.0	619	2	US-08-468-985-2
11	469.5	96.0	619	3	US-08-312-949-2
12	469.5	96.0	648	1	US-08-072-070-2
13	469.5	96.0	648	1	US-08-469-434-2
14	469.5	96.0	648	1	US-08-214-222-2
15	469.5	96.0	648	2	US-08-467-852A-2
16	469.5	96.0	648	2	US-08-468-718-2
17	469.5	96.0	648	2	US-08-247-491A-2
18	469.5	96.0	648	3	US-08-446-201-3
19	469.5	96.0	635	1	US-08-127-499A-23
20	469.5	96.0	635	1	US-08-482-847-23
21	469.5	96.0	8991	4	US-08-714-741-32
22	462.5	94.6	99	2	US-08-710-749-10
23	462.5	94.6	99	4	US-09-147-875A-11
24	462.5	94.6	204	4	US-08-529-055-51
25	457.5	93.6	288	3	US-08-312-949-4
26	457.5	93.6	288	3	US-08-446-201-4
27	447	91.4	289	1	US-08-072-070-4

28	447	91.4	289	1	US-08-469-434-4	Sequence 4, Appli
29	447	91.4	289	1	US-08-214-222-4	Sequence 4, Appli
30	447	91.4	289	1	US-08-467-852A-5	Sequence 5, Appli
31	447	91.4	289	2	US-08-468-718-4	Sequence 4, Appli
32	447	91.4	289	2	US-08-247-491A-5	Sequence 5, Appli
33	446.5	91.3	99	4	US-09-147-875A-13	Sequence 13, Appli
34	442.5	90.5	99	2	US-08-710-749-12	Sequence 12, Appli
35	442.5	90.5	195	4	US-08-529-055-71	Sequence 71, Appli
36	426.5	87.2	1231	4	US-08-714-741-41	Sequence 41, Appli
37	423.5	86.6	623	4	US-08-714-741-47	Sequence 47, Appli
38	412	84.3	100	4	US-09-147-875A-10	Sequence 10, Appli
39	407.5	83.3	99	2	US-08-710-749-17	Sequence 17, Appli
40	402.5	82.3	170	4	US-08-529-055-60	Sequence 60, Appli
41	402.5	82.3	181	4	US-08-529-055-42	Sequence 42, Appli
42	402.5	82.3	864	4	US-08-714-741-40	Sequence 40, Appli
43	399.5	81.7	99	4	US-09-147-875A-16	Sequence 16, Appli
44	399.5	81.7	188	4	US-08-529-055-59	Sequence 59, Appli
45	392.5	80.3	204	4	US-08-529-055-58	Sequence 58, Appli

ALIGNMENTS

RESULT 1
US-09-147-875A-12
; Sequence 12, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-12

Query Match 100.0%; Score 489; DB 4; Length 100;
Best Local Similarity 100.0%; Pred. No. 4e-37;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LKEIDSESEDYAKGFRAPLQSKLDKAKLSKLELSDKIDELDAEIAKLEDLQKAA 60
DB 1 LKEIDSESEDYAKGFRAPLQSKLDKAKLSKLELSDKIDELDAEIAKLEDLQKAA 60
QY 61 EENNVEDYFKGLEKTIKAAKAELEKTEADLKKA VNEPE 100
DB 61 EENNVEDYFKGLEKTIKAAKAELEKTEADLKKA VNEPE 100

RESULT 2
US-08-710-749-11
; Sequence 11, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-710-749-11

Query Match 96.0%; Score 469.5; DB 2; Length 99;
Best Local Similarity 98.0%; Pred. No. 2.2e-35;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSKDIDELDAEIAKLEDLQKAA 60
Db 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSKDIDELDAEIAKLEDLQKAA 59
Qy 61 EENNVDYFKGLEKTIAAKAELEKTEADLKKAVNEP 100
Db 60 EENNVDYFKGLEKTIAAKAELEKTEADLKKAVNEP 99

RESULT 3

US-08-529-055-61
; Sequence 61, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-61

Query Match 96.0%; Score 469.5; DB 4; Length 198;
Best Local Similarity 98.0%; Pred. No. 5e-35;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSKDIDELDAEIAKLEDLQKAA 60
Db 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSKDIDELDAEIAKLEDLQKAA 59
Qy 61 EENNVDYFKGLEKTIAAKAELEKTEADLKKAVNEP 100
Db 60 EENNVDYFKGLEKTIAAKAELEKTEADLKKAVNEP 99

RESULT 4

US-08-465-746-2
; Sequence 2, Application US/08465746
; Patent No. 5679768
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Yother, Janet L.
; APPLICANT: McDaniel, Larry S.
; TITLE OF INVENTION: EPITOPIC REGIONS OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEIN A
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/465,746
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/048,896
; FILING DATE:
; APPLICATION NUMBER: US 07/656,773
; FILING DATE: 15-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,698
; FILING DATE: 12-FEB-1992
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 521-0378
; TELEX: LUKPAT WASHINGTON
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-465-746-2

Query Match 96.0%; Score 469.5; DB 1; Length 619;
Best Local Similarity 98.0%; Pred. No. 1.9e-34;

```
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQLKAA 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQ-L-KAA 281
QY 61 EENNVEDYFKEGLEKTIKAAKAELEKTEADLKAVNEPE 100
Db 282 EENNVEDYFKEGLEKTIKAAKAELEKTEADLKAVNEPE 321

RESULT 5
US-08-214-164-2
; Sequence 2, Application US/08214164
; Patent No. 5728387
; GENERAL INFORMATION:
; APPLICANT: BRILES, DAVID E.
; APPLICANT: YOTHER, JANET L.
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 17-MAR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/656,773
; FILING DATE: 15-FEB-1991
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Berkstresser, Jerry W.
; REGISTRATION NUMBER: 22,651
; REFERENCE/DOCKET NUMBER: 6102-137
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 521-0813
; TELEX: LUKPAT WASHINGTON
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-214-164-2

Query Match 96.0%; Score 469.5; DB 1; Length 619;
Best Local Similarity 98.0%; Pred. No. 1.9e-34;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQLKAA 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQ-L-KAA 281
QY 61 EENNVEDYFKEGLEKTIKAAKAELEKTEADLKAVNEPE 100
Db 282 EENNVEDYFKEGLEKTIKAAKAELEKTEADLKAVNEPE 321

RESULT 6
US-08-467-852A-3
; Sequence 3, Application US/08467852A
; Patent No. 5856170
```

```
; GENERAL INFORMATION:
; APPLICANT: BRILES, David E.
; APPLICANT: YOTHER, Janet L.
; APPLICANT: MCDANIEL, Larry S.
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP
; STREET: 745 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10151
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,852A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOWALSKI, Thomas J.
; REGISTRATION NUMBER: 32,147
; REFERENCE/DOCKET NUMBER: 454312-2064
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-588-0800
; TELEFAX: 212-588-0500
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-467-852A-3

Query Match 96.0%; Score 469.5; DB 2; Length 619;
Best Local Similarity 98.0%; Pred. No. 1.9e-34;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQLKAA 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQ-L-KAA 281
QY 61 EENNVEDYFKEGLEKTIKAAKAELEKTEADLKAVNEPE 100
Db 282 EENNVEDYFKEGLEKTIKAAKAELEKTEADLKAVNEPE 321

RESULT 7
US-08-246-636-2
; Sequence 2, Application US/08246636
; Patent No. 5985141
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: Yother, Janet L
; APPLICANT: McDaniel, Larry S
; APPLICANT: Wu, Hong-Yin
; TITLE OF INVENTION: EPITOPIC REGIONS OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEIN A
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/246,636
;; FILING DATE: 20-MAY-1994
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/656,773
;; FILING DATE: 15-FEB-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/835,698
;; FILING DATE: 12-FEB-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/048,896
;; FILING DATE: 20-APR-1993
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 415-0810
;; TELEX: LUKPAT WASHINGTON
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 619 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-246-636-2

Query Match 96.0%; Score 469.5; DB 2; Length 619;
Best Local Similarity 98.0%; Pred. No. 1.9e-34;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLELSKDIDELDAEIAKLEDLKAA 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLELSKDIDELDAEIAKLEDLKAA 281

Qy 61 EENNVDYFKEGLEKTIAAKKAELKTEADLKAVNEPE 100
Db 282 EENNVDYFKEGLEKTIAAKKAELKTEADLKAVNEPE 321

RESULT 8
US-08-247-491A-3
; Sequence 3, Application US/08247491A
; Patent No. 5965400
; GENERAL INFORMATION:
; APPLICANT: BRILES, David E.
; APPLICANT: YOTHER, Janet L.
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP
; STREET: 745 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10151
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/247,491A
; FILING DATE: 23-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOWALSKI, Thomas J.
; REGISTRATION NUMBER: 32,147
; REFERENCE/DOCKET NUMBER: 454312-2041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-588-0800
; TELEFAX: 212-588-0500
; INFORMATION FOR SEQ ID NO: 3:

;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 619 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: n/a
;; TOPOLOGY: linear
;; MOLECULE TYPE: amino acid
US-08-247-491A-3

Query Match 96.0%; Score 469.5; DB 2; Length 619;
Best Local Similarity 98.0%; Pred. No. 1.9e-34;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLELSKDIDELDAEIAKLEDLKAA 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLELSKDIDELDAEIAKLEDLKAA 281

Qy 61 EENNVDYFKEGLEKTIAAKKAELKTEADLKAVNEPE 100
Db 282 EENNVDYFKEGLEKTIAAKKAELKTEADLKAVNEPE 321

RESULT 9
US-08-319-795-2
; Sequence 2, Application US/08319795
; Patent No. 5980909
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Yother, Janet L.
; APPLICANT: McDaniel, Larry S
; TITLE OF INVENTION: Epitopic Regions of Pneumococcal Surface
; TITLE OF INVENTION: Protein A
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheomaker and Mattcare, Ltd.
; STREET: 1203 Crystal Plaza Bldg. 1, 2001 Jefferson
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/319,795
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/246,636
; FILING DATE: 20-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/048,896
; FILING DATE: 20-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,698
; FILING DATE: 12-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/656,773
; FILING DATE: 15-FEB-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-319-795-2

Query Match 96.0%; Score 469.5; DB 2; Length 619;

Best Local Similarity 98.0%; Pred. No. 1.9e-34; Mismatches 1; Indels 0; Gaps 1;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 1 LKEIDSESEDYAKGFRAPLQSKLDKAKLSELSKIDELDAEIAKLEDLQKAA 60
Db 223 LKEIDSESEDYAKGFRAPLQSKLDKAKLSELSKIDELDAEIAKLEDLQ-KAA 281
QY 61 EENNVEDYFKGLEKTIKAAKAELEKTEADLKAVNEPE 100
Db 282 EENNVEDYFKGLEKTIKAAKAELEKTEADLKAVNEPE 321

RESULT 10
US-08-468-985-2
; Sequence 2, Application US/08468985
; Patent No. 5997882
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Yother, Janet L.
; APPLICANT: McDaniel, Larry S.
; TITLE OF INVENTION: Epitopic Regions of Pneumococcal Surface
; TITLE OF INVENTION: Protein A
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd.
; STREET: 1203 Crystal Plaza Bldg. 1, 2001 Jefferson
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,985
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/319,795
; FILING DATE:
; APPLICATION NUMBER: US 08/246,636
; FILING DATE: 20-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/048,896
; FILING DATE: 20-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,698
; FILING DATE: 12-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/656,773
; FILING DATE: 15-FEB-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-985-2

Query Match 96.0%; Score 469.5; DB 2; Length 619;
Best Local Similarity 98.0%; Pred. No. 1.9e-34;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 1 LKEIDSESEDYAKGFRAPLQSKLDKAKLSELSKIDELDAEIAKLEDLQKAA 60
Db 223 LKEIDSESEDYAKGFRAPLQSKLDKAKLSELSKIDELDAEIAKLEDLQ-KAA 281

QY 61 EENNVEDYFKGLEKTIKAAKAELEKTEADLKAVNEPE 100
Db 282 EENNVEDYFKGLEKTIKAAKAELEKTEADLKAVNEPE 321

RESULT 11
US-08-312-949-2
; Sequence 2, Application US/08312949
; Patent No. 6027734
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Wu, Hong-Yin
; TITLE OF INVENTION: MUCOSAL ADMINISTRATION OF
; TITLE OF INVENTION: PNEUMOCOCCAL ANTIGENS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,949
; FILING DATE: 30-SEP-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2049
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-312-949-2

Query Match 96.0%; Score 469.5; DB 3; Length 619;
Best Local Similarity 98.0%; Pred. No. 1.9e-34;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESEDYAKGFRAPLQSKLDKAKLSELSKIDELDAEIAKLEDLQKAA 60
Db 223 LKEIDSESEDYAKGFRAPLQSKLDKAKLSELSKIDELDAEIAKLEDLQ-KAA 281
QY 61 EENNVEDYFKGLEKTIKAAKAELEKTEADLKAVNEPE 100
Db 282 EENNVEDYFKGLEKTIKAAKAELEKTEADLKAVNEPE 321

RESULT 12
US-08-072-070-2
; Sequence 2, Application US/08072070
; Patent No. 5476929
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Yother, Janet L.
; APPLICANT: McDaniel, Larry S.
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd

STREET: Suite 1203, 2001 Jefferson Davis Highway
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202-0286
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/072,070
FILING DATE: 19930603
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/835,698
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/656,773
FILING DATE: 15-FEB-1991
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 415-0810
TELEFAX: (703) 521-0378
TELEX: LUKPAT WASHINGTON
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 648 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-072-070-2

Query Match 96.0%; Score 469.5; DB 1; Length 648;
Best Local Similarity 98.0%; Pred. No. 2e-34;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSKDIDELDAEIAKLEDLKAA 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSKDIDELDAEIAKLEDLKAA 281
Qy 61 EENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 100
Db 282 EENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 321
RESULT 13
US-08-469-434-2
Sequence 2, Application US/08469434
Patent No. 5753463
GENERAL INFORMATION:
APPLICANT: Briles, David E
APPLICANT: Yother, Janet L
APPLICANT: McDaniel, Larry S
TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Shoemaker and Mattare, Ltd
STREET: Suite 1203, 2001 Jefferson Davis Highway
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202-0286
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,434
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/072,065
FILING DATE: 03 JUNE 1993
APPLICATION NUMBER: US/07/835,698
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/656,773
FILING DATE: 15-FEB-1991
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 415-0810
TELEFAX: (703) 521-0378
TELEX: LUKPAT WASHINGTON
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 648 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-469-434-2

Query Match 96.0%; Score 469.5; DB 1; Length 648;
Best Local Similarity 98.0%; Pred. No. 2e-34;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
Qy 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSKDIDELDAEIAKLEDLKAA 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSKDIDELDAEIAKLEDLKAA 281
Qy 61 EENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 100
Db 282 EENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 321

RESULT 14
US-08-214-222-2
Sequence 2, Application US/08214222
Patent No. 5804193
GENERAL INFORMATION:
APPLICANT: Briles, David E
APPLICANT: Yother, Janet L
TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Shoemaker and Mattare, Ltd
STREET: Suite 1203, 2001 Jefferson Davis Highway
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202-0286
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/214,222
FILING DATE: 17-MAR-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/835,698
FILING DATE: 12-FEB-1992
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 415-0810
TELEFAX: (703) 521-0378
TELEX: LUKPAT WASHINGTON
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 648 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-214-222-2

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 10:12:15 ; Search time 63.8776 Seconds
(without alignments)
601.118 Million cell updates/sec

Title: US-10-674-755-12

Perfect score: 489

Sequence: 1 LKEIDSESEDIYAKGFRAP.....KKAELKTEADLKAVNEPE 100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	489	100.0	100	15	US-10-674-755-12
2	469.5	96.0	198	15	US-10-299-636-76
3	469.5	96.0	354	15	US-10-299-636-105
4	469.5	96.0	588	15	US-10-299-636-96
5	469.5	96.0	619	10	US-09-882-774-1
6	469.5	96.0	619	15	US-10-282-122A-73702
7	469.5	96.0	619	16	US-10-414-532-72
8	462.5	94.6	99	15	US-10-674-755-11
9	462.5	94.6	204	15	US-10-299-636-66
10	446.5	91.3	99	15	US-10-674-755-13
11	442.5	90.5	195	15	US-10-299-636-86

12	412	84.3	100	15	US-10-674-755-10	Sequence 10, Appl
13	402.5	82.3	170	15	US-10-299-636-75	Sequence 75, Appl
14	402.5	82.3	181	15	US-10-299-636-57	Sequence 57, Appl
15	402.5	82.3	643	15	US-10-299-636-95	Sequence 95, Appl
16	402.5	82.3	670	9	US-09-748-875-63	Sequence 63, Appl
17	402.5	82.3	670	10	US-09-298-523B-63	Sequence 63, Appl
18	402.5	82.3	690	9	US-09-748-875-61	Sequence 61, Appl
19	402.5	82.3	690	10	US-09-298-523B-61	Sequence 61, Appl
20	402.5	82.3	691	9	US-09-748-875-1	Sequence 1, Appl
21	402.5	82.3	691	10	US-09-298-523B-1	Sequence 1, Appl
22	402.5	82.3	701	9	US-09-748-875-62	Sequence 62, Appl
23	402.5	82.3	701	10	US-09-298-523B-62	Sequence 62, Appl
24	402.5	82.3	707	9	US-09-748-875-2	Sequence 2, Appl
25	402.5	82.3	707	10	US-09-298-523B-2	Sequence 2, Appl
26	402.5	82.3	711	9	US-09-748-875-3	Sequence 3, Appl
27	402.5	82.3	711	10	US-09-298-523B-3	Sequence 3, Appl
28	402.5	82.3	739	17	US-10-732-923-3294	Sequence 3294, Ap
29	402.5	82.3	929	9	US-09-748-875-60	Sequence 60, Appl
30	402.5	82.3	929	10	US-09-298-523B-60	Sequence 60, Appl
31	402.5	82.3	929	15	US-10-299-636-94	Sequence 94, Appl
32	399.5	81.7	99	15	US-10-674-755-16	Sequence 16, Appl
33	399.5	81.7	188	15	US-10-299-636-74	Sequence 74, Appl
34	392.5	80.3	204	15	US-10-299-636-73	Sequence 73, Appl
35	390.5	79.9	141	14	US-10-254-995-2	Sequence 2, Appl
36	390.5	79.9	589	9	US-09-748-875-14	Sequence 14, Appl
37	390.5	79.9	589	10	US-09-298-523B-14	Sequence 14, Appl
38	390.5	79.9	589	15	US-10-299-636-97	Sequence 97, Appl
39	387.5	79.2	99	15	US-10-674-755-15	Sequence 15, Appl
40	375.5	76.8	206	15	US-10-299-636-69	Sequence 69, Appl
41	372.5	76.2	99	15	US-10-674-755-14	Sequence 14, Appl
42	346	70.8	100	15	US-10-674-755-2	Sequence 2, Appl
43	339	69.3	100	15	US-10-674-755-3	Sequence 3, Appl
44	334.5	68.4	73	9	US-09-027-956-8	Sequence 8, Appl
45	334	68.3	194	15	US-10-299-636-79	Sequence 79, Appl

ALIGNMENTS

RESULT 1

```
US-10-674-755-12
; Sequence 12, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; PRIOR FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-12
```

Query Match 100.0%; Score 489; DB 15; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.6e-31;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKEIDSESEDIYAKGFRAPLQSKLDKQAKLSKLELSKDIDELDAETAKLEDQKAA 60

Db 1 LKEIDSESEDIYAKGFRAPLQSKLDKQAKLSKLELSKDIDELDAETAKLEDQKAA 60

QY 61 EENNVEDYFKEGLEKTIKAAKAELEKTEADLKAVNEPE 100

Db 61 EENNVEDYFKEGLEKTIKAAKAELEKTEADLKAVNEPE 100

RESULT 2

```
US-10-299-636-76
; Sequence 76, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-76

Query Match          96.0%; Score 469.5; DB 15; Length 198;
Best Local Similarity 98.0%; Pred. No. 1.9e-29;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQAKAA 60
    |||||
DB 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQAKAA 60
    |||||

QY 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 100
    |||||
DB 60 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 99
    |||||

RESULT 3
US-10-299-636-105
; Sequence 105, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 105
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-105

Query Match          96.0%; Score 469.5; DB 15; Length 354;
Best Local Similarity 98.0%; Pred. No. 3.6e-29;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQAKAA 60
    |||||
DB 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQAKAA 59
    |||||

QY 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 100
    |||||
DB 60 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 99
    |||||

RESULT 5
US-09-882-774-1
; Sequence 1, Application US/09882774
; Publication No. US20030021795A1
; GENERAL INFORMATION:
; APPLICANT: Houston, Michael E.
; APPLICANT: Hodges, Robert
; TITLE OF INVENTION: Use of Coiled-Coil Structural Scaffold to Generate
; FILE REFERENCE: 003592-007
; CURRENT APPLICATION NUMBER: US/09/882,774
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,892
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/213,387
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 619
; TYPE: PRT
US-10-299-636-96
; Sequence 96, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 96
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-96

Query Match          96.0%; Score 469.5; DB 15; Length 588;
Best Local Similarity 98.0%; Pred. No. 6.3e-29;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQAKAA 60
    |||||
DB 192 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQAKAA 250
    |||||

QY 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 100
    |||||
DB 251 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 290
    |||||

RESULT 5
US-09-882-774-1
; Sequence 1, Application US/09882774
; Publication No. US20030021795A1
; GENERAL INFORMATION:
; APPLICANT: Houston, Michael E.
; APPLICANT: Hodges, Robert
; TITLE OF INVENTION: Use of Coiled-Coil Structural Scaffold to Generate
; FILE REFERENCE: 003592-007
; CURRENT APPLICATION NUMBER: US/09/882,774
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,892
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/213,387
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 619
; TYPE: PRT
```

```
; ORGANISM: Streptococcus pneumoniae
US-09-882-774-1
Query Match          96.0%; Score 469.5; DB 10; Length 619;
Best Local Similarity 98.0%; Pred. No. 6.6e-29;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQLKAA 60
DB 223 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQ 281

QY 61 ENNNVEDYFKGLEKTIAAKKAELKTEADLKKAVNEPE 100
DB 282 ENNNVEDYFKGLEKTIAAKKAELKTEADLKKAVNEPE 321

RESULT 6
US-10-282-122A-73702
; Sequence 73702, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73702
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-282-122A-73702
Query Match          96.0%; Score 469.5; DB 15; Length 619;
Best Local Similarity 98.0%; Pred. No. 6.6e-29;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQLKAA 60
DB 223 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQ 281

; ORGANISM: Streptococcus pneumoniae
US-09-882-774-1
Query Match          96.0%; Score 469.5; DB 10; Length 619;
Best Local Similarity 98.0%; Pred. No. 6.6e-29;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 61 ENNNVEDYFKGLEKTIAAKKAELKTEADLKKAVNEPE 100
DB 282 ENNNVEDYFKGLEKTIAAKKAELKTEADLKKAVNEPE 321

RESULT 7
US-10-414-532-72
; Sequence 72, Application US/10414532
; Publication No. US2004010151A1
; GENERAL INFORMATION:
; APPLICANT: CURTISS III, ROY
; APPLICANT: KANG, HO YOUNG
; TITLE OF INVENTION: IMPROVED IMMUNOGENIC COMPOSITIONS AND VACCINES COMPRISING
; TITLE OF INVENTION: CARRIER BACTERIA THAT SECRETE ANTIGENS
; FILE REFERENCE: 56029-40437
; CURRENT APPLICATION NUMBER: US/10/414,532
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: 60/372,710
; PRIOR FILING DATE: 2002-04-16
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 72
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-414-532-72
Query Match          96.0%; Score 469.5; DB 16; Length 619;
Best Local Similarity 98.0%; Pred. No. 6.6e-29;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQLKAA 60
DB 223 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQ 281

QY 61 ENNNVEDYFKGLEKTIAAKKAELKTEADLKKAVNEPE 100
DB 282 ENNNVEDYFKGLEKTIAAKKAELKTEADLKKAVNEPE 321

RESULT 8
US-10-674-755-11
; Sequence 11, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-11
Query Match          94.6%; Score 462.5; DB 15; Length 99;
Best Local Similarity 96.0%; Pred. No. 3.2e-29;
Matches 96; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQLKAA 60
DB 1 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQ 59

QY 61 ENNNVEDYFKGLEKTIAAKKAELKTEADLKKAVNEPE 100
DB 60 ENNNVEDYFKGLEKTIAAKKAELKTEADLKKAVNEPE 99
```

```

Db          60 EENNVEDYSTEGLKTIAAKTELEKTEADLKKAVNEPE 99

RESULT 9
US-10-299-636-66
; Sequence 66, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 66
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-66

Query Match          94.6%; Score 462.5; DB 15; Length 204;
Best Local Similarity 96.0%; Pred. No. 7e-29;
Matches 96; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLELDQKKAA 60
Db 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLELDQKKAA 59

Qy 61 EENNVDYFKGGLKTIAAKAELEKTEADLKKAVNEPE 100
Db 60 EENNVDYFKGGLKTIAAKAELEKTEADLKKAVNEPE 99

RESULT 10
US-10-674-755-13
; Sequence 13, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR FILING DATE: 1999-05-24
; PRIOR APPLICATION NUMBER: US/09/147,875A
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-13

Query Match          91.3%; Score 446.5; DB 15; Length 99;
Best Local Similarity 94.0%; Pred. No. 5.8e-28;
Matches 94; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLELDQKKAA 60
Db 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLELDQKKAA 59

Qy 61 EENNVDYFKGGLKTIAAKAELEKTEADLKKAVNEPE 100
Db 60 EENNVDYFKGGLKTIAAKAELEKTEADLKKAVNEPE 100

RESULT 11
US-10-299-636-86
; Sequence 86, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-86

Query Match          90.5%; Score 442.5; DB 15; Length 195;
Best Local Similarity 93.0%; Pred. No. 2.5e-27;
Matches 93; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLELDQKKAA 60
Db 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLELDQKKAA 59

Qy 61 EENNVDYFKGGLKTIAAKAELEKTEADLKKAVNEPE 100
Db 60 EENNVEDYSTEGLKTIAAKTELEKTEADLKKAVNEPE 99

RESULT 12
US-10-674-755-10
; Sequence 10, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR FILING DATE: 1999-05-24
; PRIOR APPLICATION NUMBER: US/09/147,875A
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-10

Query Match          84.3%; Score 412; DB 15; Length 100;
Best Local Similarity 87.0%; Pred. No. 3e-25;
Matches 87; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLELDQKKAA 60
Db 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLELDQKKAA 60
```

QY 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 100
DB 61 EGNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 100

RESULT 13

US-10-299-636-75
; Sequence 75, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299, 636
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-75

Query Match 82.3%; Score 402.5; DB 15; Length 170;
Best Local Similarity 86.0%; Pred. No. 3e-24;
Matches 86; Conservative 3; Mismatches 10; Indels 1; Gaps 1;

QY 1 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLEELSDKIDELDAETAKLEDQKKAA 60
DB 1 LKEIDSESDYKGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKD-A 59

QY 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 100
DB 60 EGNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 14

US-10-299-636-57
; Sequence 57, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299, 636
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
; LENGTH: 181
; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae
US-10-299-636-57

Query Match 82.3%; Score 402.5; DB 15; Length 181;
Best Local Similarity 86.0%; Pred. No. 3.2e-24;
Matches 86; Conservative 3; Mismatches 10; Indels 1; Gaps 1;

QY 1 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLEELSDKIDELDAETAKLEDQKKAA 60
DB 1 LKEIDSESDYKGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKD-A 59
QY 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 100
DB 60 EGNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 15

US-10-299-636-95
; Sequence 95, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299, 636
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 95
; LENGTH: 643
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-95

Query Match 82.3%; Score 402.5; DB 15; Length 643;
Best Local Similarity 86.0%; Pred. No. 1.3e-23;
Matches 86; Conservative 3; Mismatches 10; Indels 1; Gaps 1;

QY 1 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLEELSDKIDELDAETAKLEDQKKAA 60
DB 245 LKEIDSESDYKGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKD-A 303

QY 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 100
DB 304 EGNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 343

Search completed: June 21, 2005, 11:18:34
Job time : 64.8775 secs

This Page Blank (uspto)

J. Bacteriol. 174: 601-609, 1992
A: Title: Structural properties and evolutionary relationships of PspA, a surface protein
A: Reference number: A41971; MUID: 92105030; PMID: 1729249
A: Accession: A41971
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-619 <YOT>
A: Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:M74122; NID:g153840; PIDN:AAA2701
A: Note: sequence extracted from NCBI backbone (NCBI:75635, NCBI:P:75636)
R: Talkington, D.F.; Crimmins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.

Query Match 24.1%; Score 118; DB 2; Length 501;
 Best Local Similarity 37.0%; Pred. No. 0.75; Mismatches 14; Indels 14; Gaps 4;
 Matches 37; Conservative 14;

QY 3 EIDSESDYAK-EGFRAPLQSKLDAQAQKLSKLELSKDIDELDAEIAKLELDQKAAE 61
 DB 125 ETTEKEVEDYKLVDSKASLDKIESAN---SOLEFNSQISLVQAANLNEQIEKLS 181

QY 62 ENNVEDYFKGLE-----KTIAAKAAELE-----KTEAD 91
 DB 182 EKNKAEQSQNEQLEFKNQNIADLIGKKAELMKLAKEED 221

RESULT 6
 H69378
 conserved hypothetical protein AF1032 - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C:Accession: H69378
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
 A:Reference number: A69250; MUID:98049343; PMID:9389475
 A:Accession: H69378
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-886 <KLE>
 A:Cross-references: UNIPROT:O92230; GB:AE001032; GB:AE000782; NID:g2689355; PIDN:AA89021
 C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 23.8%; Score 116.5; DB 2; Length 886;
 Best Local Similarity 28.1%; Pred. No. 1.6; Mismatches 22; Indels 45; Gaps 5;
 Matches 38; Conservative 22;

QY 1 LKEIDSESDYAK-EGFRAPLQSKLDAQAQKLSKLELSKDIDELDAEI----- 49
 DB 303 LRDVEKREG-DLTREA--AGIQALKAABEDNSKLEBITKRIEELERELERPEKSHRLLE 359

QY 50 -----AKLE-----DQKKAENNVNEDYFKGLEKTIKAAK 83
 DB 360 TLKPKMDRMQGIKALEENLTPDKVEMVLLSKAKEBEKEITELKK-----LIAKKS 414

QY 84 ELEKTEADLKAAVNE 98
 DB 415 SLKTRGAQLKKAVEE 429

RESULT 7
 T34418
 hypothetical protein F12F3.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T34418
 R:Fulton, B.; Wohldmann, P.
 submitted to the EMBL Data Library, July 1998
 A:Description: The sequence of C. elegans cosmid F12F3.
 A:Reference number: Z21521
 A:Accession: T34418
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-3488 <FUL>
 A:Cross-references: EMBL:U80022; PIDN:U80022; PIDN:U80022; GB:AE000023; CESP:F12F3.3
 A:Experimental source: strain Bristol N2; clone F12F3
 C:Genetics:
 A:Gene: CESP:F12F3.3
 A:Map position: 5
 A:Introns: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1

Query Match 23.7%; Score 116; DB 2; Length 3488;
 Matches 27; Conservative 22; Mismatches 29; Indels 12; Gaps 2;

Best Local Similarity 37.6%; Pred. No. 6.5; Mismatches 18; Indels 34; Gaps 9;
 Matches 47; Conservative 18;

QY 2 KEIDSESDYAK-EGFRAPLQSKLDAQAQKLSKLELSKDIDELDAEI 49
 DB 1009 KETDEKLDIAIAAKTQEADEKSKLDA-QEKIKKVSDDAARKEKELNDKL-KLESEI 1066

QY 50 A-----KLEDD-OLKKAAR-----ENNVEDYFK-----EGLEKTIKAAKAELEKTEA 90
 DB 1067 ATKKASADKLLEEQQAQKAAEVAEAAKQKQEKQKLDTEAASKAAAEKLELEK-QA 1125

QY 91 DLKKA 95
 DB 1126 QIKKA 1130

RESULT 8
 S70531
 bbbk2.11 protein precursor - Lyme disease spirochete
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C>Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
 C:Accession: S70531
 R:Akins, D.R.; Porcella, S.F.; Popova, T.G.; Shevchenko, D.; Baker, S.I.; Li, M.; Norgard
 Mol. Microbiol. 18, 507-520, 1995
 A:Title: Evidence for in vivo but not in vitro expression of a Borrelia burgdorferi outer
 A:Reference number: S70531; MUID:96342380; PMID:8748034
 A:Accession: S70531
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-233 <AKI>
 A:Cross-references: UNIPROT:Q44739; EMBL:U30617; NID:g3309515; PIDN:AA46421.1; PID:g119
 C:Superfamily: outer surface protein F ospF
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-233/Product: bbbk2.11 protein #status predicted <MAT>

Query Match 22.5%; Score 110; DB 2; Length 233;
 Best Local Similarity 31.0%; Pred. No. 1.2; Mismatches 24; Indels 24; Gaps 6;
 Matches 36; Conservative 24;

QY 6 ESESDYAK-----EGFRAPLQ-----SKLDAQAQK--LSKLEELSDKIDELDAE-- 48
 DB 37 ESEQNVKTEQEIKKQVEGFLEILETKDLSKLEDKOTKEIEKQIQELKNKIEKLSKKT 96

QY 49 ----TAKLEDQKKAAR--ENNVEDYFKGLEKTIKAAKAELEKTEADLKAAVNE 98
 DB 97 SIETTYSEBEKINKIEKLKGKLEDKPKYE-LEESLAKKKGKKAQKALQAKQKFE 151

RESULT 9
 D72230
 conserved hypothetical protein - Thermotoga maritima (strain MSB8)
 C:Species: Thermotoga maritima
 C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
 C:Accession: D72230
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
 C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
 A:Reference number: A72200; MUID:99287316; PMID:10360571
 A:Accession: D72230
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-852 <ARN>
 A:Cross-references: UNIPROT:Q9X1X1; GB:AE001806; GB:AE000512; NID:g4982196; PIDN:AA03670
 A:Experimental source: strain MSB8
 C:Genetics:
 A:Gene: TM1636
 C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 22.5%; Score 110; DB 2; Length 852;
 Best Local Similarity 30.0%; Pred. No. 4.1; Mismatches 27; Indels 12; Gaps 2;

RESULT 12

F75216

hypothetical protein PAB2181 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C:Accession: F75216

R:anonymous, Genoscope

A:Submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure

A:Reference number: A75001

A:Accession: F75216

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-281 <RAW>

A:Cross-references: UNIPROT:Q9V217; GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB4918

C:Genetics:

A:Gene: PAB2181

Query Match 22.2%; Score 108.5; DB 2; Length 281;

Best Local Similarity 27.9%; Pred. No. 1.8;

Matches 29; Conservative 27; Mismatches 39; Indels 9; Gaps 2;

Qy 2 KEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLELSKDIDELDAEIAKLEDOLK---- 57

Db 163 KEIELKGKVEKLEQKEKLEKLEKSEVKLMYEAKAKAELEAKLREYEEKVGRREE 222

Qy 58 ---KAAEENNVEDYFKGLEKXTIAAKAELEKTEADLKAVNE 98

Db 223 LERKVSLESLSENEY--ETKVSLEKKEELENKVKLEEEVVK 264

RESULT 13

S70532

outer surface protein F precursor - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004

C:Accession: S70532

R:Akins, D.R.; Porcellia, S.F.; Popova, T.G.; Shevchenko, D.; Baker, S.I.; Li, M.; Norgard

Mol. Microbiol. 18, 507-520, 1995

A:Title: Evidence for in vivo but not in vitro expression of a Borrelia burgdorferi outer

A:Reference number: S70531; MUID:96342380; PMID:8748034

A:Accession: S70532

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-229 <AKI>

A:Cross-references: UNIPROT:Q44735; EMBL:U19754; NID:g318660; PIDN:AAC26147.1; PID:g896

C:Genetics:

A:Gene: ospF

C:Superfamily: outer surface protein F ospF

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-229/Product: outer surface protein F #status predicted <MAT>

Query Match 22.0%; Score 107.5; DB 2; Length 229;

Best Local Similarity 31.1%; Pred. No. 1.7;

Matches 37; Conservative 27; Mismatches 32; Indels 23; Gaps 7;

Qy 1 LKEIDSE-----SEDIYAK---EGFRAPLQ-----SKLDAKQAKLSKLELSKDIDELDAE 48

Db 33 VQDLESSEQNKKTEQIKKQVGVFLEILTKDLNLDLTKIEI-KRIQELKEKIEKLEAK 91

Qy 49 IAKL-----EDOLKAAEE-----NNVEDYFKGLEKXTIAAKAELEKTEADLKAVNE 98

Db 92 KTSLKTYSEYEKLEKQIKLEKGLKADLEKLGK-GLSDSLKKKKKEERKKALEDKAKKFFEE 149

RESULT 14

T05409

hypothetical protein F10M6.170 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C:Accession: T05409
R:Bevan, M.; Weichselgartner, M.; Fartmann, B.; Granderath, K.; Dauner, D.; Herzl, A.; N
submitted to the Protein Sequence Database, February 1998
A:Reference number: Z15414
A:Accession: T05409
A:Molecule type: DNA
A:Residues: 1-764 <BEV>
A:Cross-references: UNIPROT:O49371; EMBL:AL021811
A:Experimental source: cultivar Columbia; BAC clone F10M6
C:Genetics:
A:Map position: 4
A:Note: F10M6.170

Query Match 22.0%; Score 107.5; DB 2; Length 764;
Best Local Similarity 29.8%; Pred. No. 5.4;
Matches 36; Conservative 23; Mismatches 31; Indels 31; Gaps 4;
QY 2 KEIDSESDYAKGFRAPLQSKDAQAKLSKL-EELSDKIDELDAETAKLEDQ----- 55
DB 163 REIEELKHKLRDERERAAQLQSLTUKKEELEKQREIARSKESVMAISEFESKQLLS 222
QY 56 -----LKKAAEENNVEDYFKGLEKTIAAKKALEK---TEADLKKAVN 97
DB 223 KANEVVRQEGEYIALQRALEK-----EELEISKATKKLEQKLRTEANLKKQTE 275
QY 98 E 98
DB 276 E 276

RESULT 15
T30845
probable DNA repair protein RAD50 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30845
R:Kim, K.K.; Daud, A.I.; Wong, S.C.; Pajak, L.; Tsai, S.C.; Wang, H.; Henzel, W.J.; Fiel
J. Biol. Chem. 271, 29255-29284, 1996
A:Title: Mouse RAD50 has limited epitopic homology to p53 and is expressed in the adult
A:Reference number: Z20899; MUID:97067183; PMID:8910585
A:Accession: T30845
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1312 <KIM>
A:Cross-references: UNIPROT:P70388; EMBL:U66887; NID:g1575574; PID:g1575575; PIDN:AAC528
C:Genetics:
A:Gene: RAD50
A:Map position: 11
C:Superfamily: RAD50 protein
C:Keywords: DNA repair

Query Match 21.5%; Score 105; DB 2; Length 1312;
Best Local Similarity 35.1%; Pred. No. 13;
Matches 27; Conservative 20; Mismatches 20; Indels 10; Gaps 2;
QY 22 QSKLDKAKLSKLELSKDIDELDAETAKLEDQKAAEENNVEDYFKGLEKTIAAK 81
DB 454 QSELHVRSELQLEGSDRILELDQELTKAERLSK-AEKNSSIE-----TLKAE 503
QY 82 KAELEKTEADLKKAUNE 98
DB 504 VMSLQNEKADLDKSLRK 520

Search completed: June 21, 2005, 10:11:58
Job time : 11 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:54:24 ; Search time 61.9388 Seconds
(without alignments)
826.751 Million cell updates/sec

Title: US-10-674-755-12

Perfect score: 489

Sequence: 1 LKEIDSESEDYAKGFRAP.....KKAELEKTEADLKAVNEPE 100

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	469.5	96.0	619	2 Q54972	Q54972 streptococc
2	469.5	96.0	619	2 Q8DR10	Q8DR10 streptococc
3	462.5	94.6	417	2 Q9LAY3	Q9LAY3 streptococc
4	447.5	91.5	415	2 Q9LAY1	Q9LAY1 streptococc
5	402.5	82.3	739	2 Q9ROT4	Q9ROT4 streptococc
6	402.5	82.3	820	2 Q9ROT1	Q9ROT1 streptococc
7	402.5	82.3	929	2 Q9KK19	Q9KK19 streptococc
8	402.5	82.3	929	2 Q9ZAY5	Q9ZAY5 streptococc
9	389.5	79.7	437	2 Q9LAY4	Q9LAY4 streptococc
10	387.5	79.2	395	2 Q9LAY2	Q9LAY2 streptococc
11	387.5	79.2	408	2 Q9LAY0	Q9LAY0 streptococc
12	383.5	78.4	99	2 Q8KOK4	Q8KOK4 streptococc
13	383.5	78.4	249	2 Q9LS75	Q9LS75 streptococc
14	376.5	77.0	224	2 Q8GNS8	Q8GNS8 streptococc
15	370.5	75.8	426	2 Q9LAY5	Q9LAY5 streptococc
16	368	75.3	869	2 Q9KK27	Q9KK27 streptococc
17	357	73.0	225	2 Q9LS91	Q9LS91 streptococc
18	351	71.8	222	2 Q9LS77	Q9LS77 streptococc
19	351	71.8	262	2 Q9LS76	Q9LS76 streptococc
20	351	71.8	415	2 Q9LAY7	Q9LAY7 streptococc
21	347	71.0	246	2 Q9LS78	Q9LS78 streptococc
22	344	70.3	416	2 Q9LAY8	Q9LAY8 streptococc
23	342	69.9	255	2 Q9LS81	Q9LS81 streptococc
24	342	69.9	255	2 Q9LS86	Q9LS86 streptococc
25	340	69.5	406	2 Q9LAZ0	Q9LAZ0 streptococc
26	335	68.5	393	2 Q9LAZ3	Q9LAZ3 streptococc
27	334	68.3	394	2 Q9LAY6	Q9LAY6 streptococc
28	334	68.3	395	2 Q9LAZ1	Q9LAZ1 streptococc
29	329	67.3	340	2 Q8KOK5	Q8KOK5 streptococc
30	328	67.1	194	2 Q9LSB5	Q9LSB5 streptococc
31	328	67.1	218	2 Q6UEB2	Q6UEB2 streptococc

32 328 67.1 233 2 Q9LS68 streptococc
33 328 67.1 236 2 Q9LS69 streptococc
34 328 67.1 243 2 Q9LS64 streptococc
35 328 67.1 243 2 Q9LS67 streptococc
36 328 67.1 244 2 Q9LS65 streptococc
37 328 67.1 247 2 Q9LS66 streptococc
38 328 67.1 249 2 Q9LS70 streptococc
39 328 67.1 254 2 Q9LS63 streptococc
40 328 67.1 401 2 Q9LAZ2 streptococc
41 326 66.7 207 2 Q8GNS9 streptococc
42 320 65.4 237 2 Q9LS92 streptococc
43 320 65.4 395 2 Q9LAY9 streptococc
44 216 44.2 653 2 Q34097 streptococc
45 200 40.9 246 2 Q9LSB4 streptococc

ALIGNMENTS

RESULT 1
Q54972 PRELIMINARY; PRT; 619 AA.
AC Q54972;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pneumococcal surface protein A precursor.
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92105030; PubMed=1729249;
RA Yother J., Briles D.E.;
RT "Structural properties and evolutionary relationships of PspA, a surface protein of Streptococcus pneumoniae, as revealed by sequence analysis."
RL J. Bacteriol. 174:601-609(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Yother J., Briles D.E.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M74122; AAA27018.1; -.
DR PIR; A41971; A41971.
DR PIR; A97887; A97887.
DR HSSP; P06653; LHCX.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam; PF01473; CW_binding_1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_3.
KW Signal.
FT SIGNAL. 1 31 Potential.
FT CHAIN 32 619 pneumococcal surface protein A.
SQ SEQUENCE 619 AA; 68605 MW; 5AA8BDB40C2841CA CRC64;

Query Match 96.0%; Score 469.5; DB 2; Length 619;
Best Local Similarity 98.0%; Pred. No. 8.6e-21;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQKAA 60
|||||
Db 223 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQKAA 281
|||||

QY 61 EENNVEDYFKGLEKTIKAAKAELEKTEADLKAVNEPE 100
|||||

Db 282 EENNVEDYFKGLEKTIKAAKAELEKTEADLKAVNEPE 321
|||||

RESULT 2

Q8DR10 PRELIMINARY; PRT; 619 AA.
ID Q8DR10

AC	QBRD10;
AD	01-MAR-2003 (TrEMBLrel. 23, Created)
AE	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
AF	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
AG	Surface protein pspA.
AH	Name=pspA; OrderedLocusNames=spr0121;
AI	Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
AJ	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
AK	Streptococcus.
AL	NCBI_TaxID=171101;
AM	[1]
AN	SEQUENCE FROM N.A.
AO	MEDLINE=21429245; PubMed=11544234;
AP	DOI=10.1128/JB.183.19.5709-5717.2001;
AQ	RX Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
AR	DeHoff B.S., Etrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
AS	Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
AT	LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
AV	McAhren S.M., McGarvey M., McEaster K., Mundy C.W., Niclas T.I.,
AW	Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rockey P.,
AX	Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
AY	Zook C.A., Baltz R.H., Jaskunas S.R., Rostock P.R. Jr., Skatrud P.L.,
AZ	Glass J.I.;
BA	"genome of the bacterium Streptococcus pneumoniae strain R6.";
BB	J. Bacteriol. 183:5709-5717(2001).
BC	ENBL; AE008396; AAK98925.1; -.
BD	PIR; A41971; A41971.
BE	PIR; A97887; A97887.
BF	HSSP; P06653; IHXX.
BG	InterPro; IPR002479; CW binding.
BH	InterPro; IPR002345; Lipocalin.
BI	Pfam; PF01473; CW binding_1; 10.
BJ	PROSITE; PS00213; LIPOCALIN; UNKNOWN_3.
BK	Complete proteome.
BL	SEQUENCE 619 AA; 68605 MW; 5AA8EBD40C2841CA CRC64;
BM	
BN	
BO	
BP	
BQ	
BR	
BS	
BT	
BU	
BV	
BW	
BX	
BY	
BZ	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	</


```

RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, PspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL: AF068647; AAF13457.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR002479; CW_binding.
DR InterPro: IPR005877; Gpos_Ysirk.
DR InterPro: IPR009053; Prefoldin.
DR InterPro: IPR007756; RICH.
DR Pfam: PF01473; CW_binding_1; 1.
DR Pfam: PF05062; RICH; 2.
DR Pfam: PF04650; Ysirk_signal; 1.
DR TIGRFAMs: TIGR01168; Ysirk_signal; 1.
KW Hypothetical protein.
FT NON_TER 739
SQ SEQUENCE 739 AA; 83960 MW; 7EE2F2F676ABF989 CRC64;

Query Match 82.3%; Score 402.5; DB 2; Length 739;
Best Local Similarity 86.0%; Pred. No. 1.1e-16;
Matches 86; Conservative 3; Mismatches 10; Indels 1; Gaps 1;

QY 1 LKEIDSESDYAKGFRAPLOSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQLKAA 60
DB 537 LKEIDSESDYAKGFRAPLOSKLDTKKAKLKLEELSDKIDELDAETAKLEVQLKD-A 595
QY 61 EENNVEDYFKEGLEKTTAAKKAELKTEADLKAVNEPE 100
DB 596 EGNNVVEAYFKEGLEKTTAAKKAELKAEADLKAVDEPE 635

RESULT 6
Q9RQT1 PRELIMINARY; PRT; 820 AA.
AC Q9RQT1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 26, Last annotation update)
DE Hypothetical protein pspC (fragment).
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=BG9163;
RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, PspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL: AF068650; AAF13460.1; -.
DR HSSP: P04268; 1IC2.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR002479; CW_binding.
DR InterPro: IPR005877; Gpos_Ysirk.
DR InterPro: IPR009053; Prefoldin.
DR InterPro: IPR007756; RICH.
DR Pfam: PF01473; CW_binding_1; 1.
DR Pfam: PF05062; RICH; 2.
DR Pfam: PF04650; Ysirk_signal; 1.
DR TIGRFAMs: TIGR01168; Ysirk_signal; 1.
KW Hypothetical protein.
FT NON_TER 820
SQ SEQUENCE 820 AA; 91752 MW; 33C095849ABB0942 CRC64;

Query Match 82.3%; Score 402.5; DB 2; Length 820;
Best Local Similarity 86.0%; Pred. No. 1.2e-16;
Matches 86; Conservative 3; Mismatches 10; Indels 1; Gaps 1;

QY 1 LKEIDSESDYAKGFRAPLOSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQLKAA 60
DB 537 LKEIDSESDYAKGFRAPLOSKLDTKKAKLKLEELSDKIDELDAETAKLEVQLKD-A 595
QY 61 EENNVEDYFKEGLEKTTAAKKAELKTEADLKAVNEPE 100
DB 596 EGNNVVEAYFKEGLEKTTAAKKAELKAEADLKAVDEPE 635

RESULT 7
Q9KK19 PRELIMINARY; PRT; 929 AA.
AC Q9KK19;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Surface protein PspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=srf10;
RX MEDLINE=21888621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae.";
RL Gene 284:63-71(2002).
DR EMBL: AF154037; AAF73809.1; -.
DR HSSP: P06653; 1H8G.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR002479; CW_binding.
DR InterPro: IPR005877; Gpos_Ysirk.
DR InterPro: IPR009053; Prefoldin.
DR InterPro: IPR007756; RICH.
DR Pfam: PF01473; CW_binding_1; 1.
DR Pfam: PF05062; RICH; 2.
DR Pfam: PF04650; Ysirk_signal; 1.
DR TIGRFAMs: TIGR01168; Ysirk_signal; 1.
SQ SEQUENCE 929 AA; 105003 MW; 2DC8293302FAFA64 CRC64;

Query Match 82.3%; Score 402.5; DB 2; Length 929;
Best Local Similarity 86.0%; Pred. No. 1.3e-16;
Matches 86; Conservative 3; Mismatches 10; Indels 1; Gaps 1;

QY 1 LKEIDSESDYAKGFRAPLOSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQLKAA 60
DB 530 LKEIDSESDYAKGFRAPLOSKLDTKKAKLKLEELSDKIDELDAETAKLEVQLKD-A 588
QY 61 EENNVEDYFKEGLEKTTAAKKAELKTEADLKAVNEPE 100
DB 589 EGNNVVEAYFKEGLEKTTAAKKAELKAEADLKAVDEPE 628

RESULT 8
Q9ZAY5 PRELIMINARY; PRT; 929 AA.
AC Q9ZAY5;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Surface protein C.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=BF6796;

```

```

RX MEDLINE=20038319; PubMed=10569772;
RA Brooke-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, pspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; U72655; AAD00184.1; -.
DR HSP; P06653; 1HCX.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_YSRK.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; YSRK signal; 1.
DR TIGRFAMs; TIGR01168; YSRK; signal; 1.
SQ SEQUENCE 929 AA; 104991 MW; 2DC8293302PFB081 CRC64;

Query Match 82.3%; Score 402.5; DB 2; Length 929;
Best Local Similarity 86.0%; Pred. No. 1.3e-16;
Matches 86; Conservative 3; Mismatches 10; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLELSKIDELDAEIAKLELDQKAA 60
Db 530 LKEIDSESDYAKGFRAPLQSKLDTKAKLSKLELSKIDELDAEIAKLEVLQKD-A 588

Qy 61 EENNVDYFKGLEKTIKAELEKTEADLKKAVNEP 100
Db 589 EGNNVAYFKGLEKTTAEKKALEKAEADLKKAVDEP 628

RESULT 9
Q9LAY4
ID Q9LAY4 PRELIMINARY; PRT; 437 AA.
AC Q9LAY4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E134;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071813; AAF27709.1; -.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin alc DH like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 437
SQ SEQUENCE 437 AA; 48071 MW; F6EPD2D13E08CD8 CRC64;

Query Match 79.7%; Score 389.5; DB 2; Length 437;
Best Local Similarity 84.0%; Pred. No. 4.2e-16;
Matches 84; Conservative 3; Mismatches 12; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLELSKIDELDAEIAKLELDQKAA 60
Db 235 LKEIDSESDYKXKEGLRAPLQSKLDTKAKLSKLELSKIDELDAEIAKHVQLKD-A 293

Qy 61 EENNVDYFKGLEKTIKAELEKTEADLKKAVNEP 100
Db 294 EGNNVAYFKGLEKTTAEKKALEKAEADLKKAVDEP 333

us-10-674-755-12.rup

```

```

RESULT 10
Q9LAY2
ID Q9LAY2 PRELIMINARY; PRT; 395 AA.
AC Q9LAY2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EP6796;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071813; AAF27709.1; -.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin alc DH like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 395
SQ SEQUENCE 395 AA; 42963 MW; 586EF956BCBCCLE CRC64;

Query Match 79.2%; Score 387.5; DB 2; Length 395;
Best Local Similarity 83.0%; Pred. No. 5.1e-16;
Matches 83; Conservative 7; Mismatches 9; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLELSKIDELDAEIAKLELDQKAA 60
Db 225 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLELSKIDELDAEIAEVLQKD-A 283

Qy 61 EENNVDYFKGLEKTIKAELEKTEADLKKAVNEP 100
Db 284 EGNNVAYFKGLEKTTAEKKALEKAEADLKKAVDEP 323

RESULT 11
Q9LAY0
ID Q9LAY0 PRELIMINARY; PRT; 408 AA.
AC Q9LAY0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9163;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071815; AAF27711.1; -.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin alc DH like.
FT NON TER 408
SQ SEQUENCE 408 AA; 44254 MW; 4F64D874217297EF CRC64;

Query Match 79.2%; Score 387.5; DB 2; Length 408;
Best Local Similarity 83.0%; Pred. No. 5.3e-16;

```


DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBL5;
RX MEDLINE=20448953; PubMed=10992499;
RK DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae."
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071810; AAP27706.1; -.
DR HSSP; P00192; IM6T.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 426
SQ SEQUENCE 426 AA; 46534 MW; 81AA11348CBE6634 CRC64;

Query Match 75.8%; Score 370.5; DB 2; Length 426;
Best Local Similarity 79.0%; Pred. No. 5.8e-15;
Matches 79; Conservative 8; Mismatches 12; Indels 1; Gaps 1;

Qy 1 LKEIDSESEDYAKEGFRAPLQSKLDKAKQAKLSKLELSKDIDELDAEIAKLEDLKAA 60
Db 215 LKIDNESDSEDYVKEGURAPLQSELDTKAKLLKLELSGKIIEELDAEIAELEVQLKD-A 273

Qy 61 EENNVEDYFKGLEKTIKAAKAELEKTEADLKAVNEPE 100
Db 274 EGNNVVAYPKGLEKTIKAAKAELEKTEADLKAVNEPE 313

Search completed: June 21, 2005, 10:22:10
Job time : 61.9388 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:50:39 ; Search time 73.8459 Seconds
(without alignments)
518.502 Million cell updates/sec

Title: US-10-674-755-13

Perfect score: 486

Sequence: 1 LKEIDESESDYAKGFRAP.....KKTELEKTEADLKXAVNEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	482	99.2	195	2	AAW14591 Streptoco
2	482	99.2	195	7	ABW02625 Wu2c pneu
3	482	99.2	8991	6	ABU08487 S. pneumo
4	453	93.2	198	7	ABW02615 Rx1c pneu
5	453	93.2	315	2	AAW04375 Streptoco
6	453	93.2	619	2	AAW63437 Pneumococ
7	453	93.2	619	2	AAW87598 Pneumococ
8	453	93.2	619	2	AAW86911 Pneumococ
9	453	93.2	619	2	AAW41838 Streptoco
10	453	93.2	619	5	AAE18782 S. pneumo
11	453	93.2	619	6	ABU45778 Protein e
12	453	93.2	619	8	ADOS2126 Streptoco
13	453	93.2	648	2	AAW70336 Pneumococ
14	453	93.2	648	2	AAW62274 Streptoco
15	453	93.2	648	2	AAW41837 Streptoco
16	453	93.2	648	2	AAW87879 A pneumoc
17	453	93.2	653	2	AAW92456 S. pneumo
18	453	93.2	684	2	AAW73912 Streptoco
19	450	92.6	198	2	AAW14581 Streptoco
20	446	91.8	204	2	AAW14571 Streptoco
21	446	91.8	204	7	ABW02605 Ef1019c p
22	433	89.1	653	2	AAW27150 PspA frag
23	430.5	88.6	289	2	AAW62276 Streptoco
24	430.5	88.6	289	2	AAW41840 Streptoco
25	430.5	88.6	289	2	AAW87910 Protein s

26	430.5	88.6	289	2	AAW92458 S. pneumo
27	407	83.7	623	6	ABU08494 Fragment
28	394	81.1	1231	6	ABU08490 Fragment
29	385	79.2	170	7	ABW02614 Rct135c p
30	385	79.2	181	7	ABW02596 0922134c
31	385	79.2	865	6	ABU08489 S. pneumo
32	385	79.2	929	2	AAW14593 Streptoco
33	385	79.2	929	2	AAW43384 S. pneumo
34	382	78.6	188	2	AAW14580 Streptoco
35	382	78.6	188	7	ABW02613 Rct139c p
36	375	77.2	204	2	AAW14578 Streptoco
37	375	77.2	204	7	ABW02612 Rct123c p
38	373	76.7	588	6	ABU08491 Coiled co
39	373	76.7	588	2	AAW43392 PspC alph
40	370.5	76.2	180	2	AAW14562 Streptoco
41	367.5	75.6	187	2	AAW14579 Streptoco
42	358	73.7	206	2	AAW14574 Streptoco
43	358	73.7	206	7	ABW02608 Db15c pne
44	332.5	68.4	550	8	ADK48356 Streptoco
45	332.5	68.4	550	8	ADR95223 Novel S.

ALIGNMENTS

RESULT 1

AAW14591

ID AAW14591 standard; protein; 195 AA.

XX AAW14591;

XX 17-OCT-2003 (revised)

DT 28-OCT-1997 (first entry)

XX

DE Streptococcus pneumoniae PspA central region.

XX

PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
bacteraemia; pneumonia.

XX

OS Streptococcus pneumoniae; strain Wu2.

XX

PN WO709994-A1.

XX

PD 20-MAR-1997.

XX

PF 16-SEP-1996; 96WO-US014819.

XX

PR 15-SEP-1995; 95US-00529055.

XX

PA (UABR-) UAB RES FOUND.

XX

PI Briles DB, McDaniel LS, Swiatlo E, Yother J, Crain MJ;

XX

PI Hollingshead S, Tart R, Brooks-Walter A;

XX

DR WPI; 1997-202002/18.

XX

PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used
in vaccines for protecting animals against S.pneumoniae infection.

XX

PS Example 6; Fig 13; 296pp; English.

XX

CC This sequence shows the central portion, including the C-terminus of the
alpha-helix region and some of the proline-rich region, of pneumococcal

CC

CC surface protein A (PspA) of Streptococcus pneumoniae strain Wu2.

CC

CC Comparison of the N-terminal and central regions (AAW14533-57 and
AAW14562-91) of PspA polypeptides from different pneumococcal strains can

CC

CC be used to divide the strains into several families based on sequence
homologies. PspA polypeptides, or fragments of them, can be used in

CC

CC vaccines to protect animals against S. pneumoniae infection and hence for
the prevention of diseases such as otitis media, meningitis, bacteraemia

CC

CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
region and the immediate 5' tip of the coding sequence are likely to be

CC

CC the critical sequences for predicting PspA cross-reactions and vaccine

```
CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 195 AA;

Query Match          99.2%; Score 482; DB 2; Length 195;
Best Local Similarity 99.0%; Pred. No. 1.4e-35;
Matches 98; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDSESEDYAKGFRAPLHSLDKAKOAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
    |||||
Db 1 LKEIDSESEDYAKGFRAPLHSLDKAKOAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
    |||||

Qy 61 ENNVEDYSTEGLEKTIAAKKTELEKTEADLKKA VNEPE 99
    |||||
Db 61 ENNVEDYSTEGLEKTIAAKKTELEKTEADLKKA VNEPE 99
    |||||

RESULT 2
ABW02625
ID ABW02625 standard; protein; 195 AA.
XX
AC ABW02625;
XX
XX
XX 12-FEB-2004 (first entry)
DT
DE Wu2c pneumococcal surface protein A (PspA) central region.
XX
XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
KW immunological; gene therapy; immunostimulant.
XX
XX Unidentified.
XX
XX US6592876-B1.
XX
XX 15-JUL-2003.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX 20-APR-1993; 93US-00048896.
XX
XX 06-JUN-1995; 95US-00465746.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX
XX WPI; 2003-862841/80.
XX
XX Immunological composition for obtaining expression products used for
PT detecting the presence of Streptococcus pneumoniae or its strain,
PT comprises at least two different full length isolated gene encoding
PT pneumococcal surface protein A.
XX
XX
XX Example 6; SEQ ID NO 71; 121pp; English.
XX
XX The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspA) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antigenic,
CC immunological or vaccine compositions, for eliciting antibodies, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Wu2c pneumococcal
CC surface protein A (PspA) central region. This sequence is used in the
CC exemplification of the invention
XX
XX Sequence 195 AA;

Query Match          99.2%; Score 482; DB 7; Length 195;
Best Local Similarity 99.0%; Pred. No. 1.4e-35;

Qy 1 LKEIDSESEDYAKGFRAPLHSLDKAKOAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
    |||||
Db 8797 LKEIDSESEDYAKGFRAPLHSLDKAKOAKLSKLELSKIDELDAEIAKLEDLQKAVE 8856
    |||||

Matches 98; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Query Match          99.2%; Score 482; DB 6; Length 8991;
Best Local Similarity 99.0%; Pred. No. 1.5e-33;
Matches 98; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDSESEDYAKGFRAPLHSLDKAKOAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
    |||||
Db 8797 LKEIDSESEDYAKGFRAPLHSLDKAKOAKLSKLELSKIDELDAEIAKLEDLQKAVE 8856
    |||||
```

QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKAVNEPE 99
 DB 8857 ENNVEDYSTEGLEKTIAAKTELEKTEADLKAVNEPE 8895

RESULT 4

ABW02615
 ID ABW02615 standard; protein; 198 AA.

XX AC ABW02615;
 XX DT 12-FEB-2004 (first entry)
 XX DE RxlC pneumococcal surface protein A (PspA) central region.
 XX KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
 XX KW immunological; gene therapy; immunostimulant.
 XX OS Unidentified.
 XX PN US6592876-B1.
 XX PD 15-JUL-2003.
 XX PF 15-SEP-1995; 95US-00529055.
 XX PR 20-APR-1993; 93US-00048896.
 XX PR 06-JUN-1995; 95US-00465746.
 XX PA (UABR-) UAB RES FOUND.

XX PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
 XX DR WPI; 2003-862841/80.

XX Immunological composition for obtaining expression products used for
 PT detecting the presence of Streptococcus pneumoniae or its strain,
 PT comprises at least two different full length isolated gene encoding
 PT pneumococcal surface protein A.

XX Example 6; SEQ ID NO 61; 121pp; English.

XX The present invention relates to an immunological composition comprising
 CC at least 2 different full length isolated genes encoding pneumococcal
 CC surface protein A (PspAs) from different groups based on restriction
 CC fragment polymorphism analysis. The invention is useful for obtaining
 CC expression products by recombinant techniques to detect, determine,
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its
 CC strain. The expression product is useful for preparing antigenic,
 CC immunological or vaccine compositions, for eliciting antibodies, an
 CC immunological response (other than or additional to antibodies) or a
 CC protective response (including antibody or other immunological response
 CC by administering compositions to a host). The invention is also useful as
 CC vaccines and in gene therapy. The present sequence is RxlC pneumococcal
 CC surface protein A (PspA) central region. This sequence is used in the
 CC exemplification of the invention

XX SQ Sequence 198 AA;

Query Match 93.2%; Score 453; DB 7; Length 198;
 Best Local Similarity 93.9%; Pred. No. 6.1e-33;
 Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKAKLSELSKDIDELDAETAKLEDQKAVE 60
 DB 1 LKEIDSESDYAKGFRAPLHSLDKAKLSELSKDIDELDAETAKLEDQKAAE 60

QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKAVNEPE 99

DB 61 ENNVEDYFKEGLEKTIAAKAELEKTEADLKAVNEPE 99

RESULT 5

AAV04375

ID AAY04375 standard; protein; 315 AA.

XX AC AAY04375;

XX DT 23-JUN-1999 (first entry)

XX DE Streptococcus pneumoniae PspA protein sequence.

XX KW Streptococcus pneumoniae; pspA; pneumococcal; surface protein; vaccine;
 XX KW immunological; infection.

XX OS Streptococcus pneumoniae.

XX OS Synthetic.

XX PN WO9914333-A2.

XX PD 25-MAR-1999.

XX PF 18-SEP-1998; 98WO-US019740.

XX PR 18-SEP-1997; 97US-00932982.

XX PA (INMR) PASTEUR MERIEUX CONNAUGHT.

XX PI Becker R, Gray M, Pyle D;

XX WPI; 1999-229537/19.

XX DR N-PSDB; AAX33124.

XX PT DNA encoding PspA molecule with modified internal translational
 PT initiation sites.

XX PS Disclosure; Page; 36pp; English.

XX The present sequence represents a pneumococcal surface protein A (PspA)
 CC molecule where internal naturally occurring translational initiation
 CC sites have been modified or eliminated so that expression of the DNA
 CC sequence results in a single form of PspA. The PspA nucleotide sequence
 CC can be used to transform a unicellular host to produce the PspA protein.
 CC The PspA protein can be used in an immunological composition for treating
 CC or preventing S. pneumoniae infection especially in a child. Antibodies
 CC to the PspA protein can also be used to treat S. pneumoniae infection.
 CC The immunogenic peptides are designed to confer broad protection against
 CC diverse pneumococcal strains. N.B. The present sequence is not given in
 CC the specification but is encoded by the sequence given in AAX33124

XX SQ Sequence 315 AA;

Query Match 93.2%; Score 453; DB 2; Length 315;
 Best Local Similarity 93.9%; Pred. No. 1.1e-32;
 Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKAKLSELSKDIDELDAETAKLEDQKAVE 60
 DB 193 LKEIDSESDYAKGFRAPLHSLDKAKLSELSKDIDELDAETAKLEDQKAAE 252

QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKAVNEPE 99

DB 253 ENNVEDYFKEGLEKTIAAKAELEKTEADLKAVNEPE 291

RESULT 6

AAR63437

ID AAR63437 standard; protein; 619 AA.

XX AC AAR63437;

XX DT 09-SEP-2004 (revised)

DT 16-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 19-JUL-1995 (first entry)

XX

```

DE Pneumococcal surface protein A from S.pneumoniae Rx1.
XX
XX Pneumococcal surface protein A; PspA; Streptococcus; PCR; pneumococcal;
KW primer; protection-eliciting epitope; epitope; vaccine; amplify.
XX
XX Streptococcus pneumoniae.
OS
OS Unidentified.
XX
XX Key Location/Qualifiers
FH 192..260
FT Protein /note= "protein fragment of Claim 1"
FT
XX
XX EP622081-A2.
XX
XX 02-NOV-1994.
XX
XX 19-APR-1994; 94EP-00302767.
XX
XX 20-APR-1993; 93US-00048896.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, Yother JL, McDaniel LS;
XX WPI; 1994-359522/45.
DR N-PSDB; AAQ78131.
XX
XX regions of Pneumococcal surface protein A - derived from the Rx1 PspA
PT strain, for the preparation of cross-reactive vaccines for the prevention
PT of pneumococcal infections.
XX
XX Disclosure; Page 13-16; 26pp; English.
XX
XX The amino acid sequence of the novel Pneumococcal surface protein A
CC (PspA) from Streptococcus pneumoniae strain Rx1. The gene was PCR
CC amplified from S.pneumoniae genomic DNA using the primers AAQ78132-5. The
CC gene was used to derive truncated peptide fragments containing protection
CC -eliciting epitopes for use in vaccines against pneumococcal diseases.
CC The epitopic fragments are derived from amino acids 192-260 and
CC optionally contain a further 25 a.a. residues at both the N- and C-
CC terminal regions of the peptide. The epitopic peptide fragments may be
CC derived from different strains of S.pneumoniae and are homologous to the
CC Rx1 strain epitope. (Updated on 25-MAR-2003 to correct PN field.)
CC (Updated on 16-OCT-2003 to standardise OS field)
CC
CC Revised record issued on 09-SEP-2004 : Correction to feature table key
XX
XX Sequence 619 AA;
SQ
Query Match 93.2%; Score 453; DB 2; Length 619;
Best Local Similarity 93.9%; Pred. No. 2.4e-32;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 1 LKEIDSESDYAKGFRAPHSKLDKAKQKLSKLELSKIDELDAEIAKLELDQKAVE 60
Db 223 LKEIDSESDYAKGFRAPHSKLDKAKQKLSKLELSKIDELDAEIAKLELDQKAAE 282
Qy 61 ENNVVDYSTGLEKTTAAKKTLEKTEADLKKA VNEPE 99
Db 283 ENNVVDYFREGLEKTTAAKKAELKTEADLKKA VNEPE 321
RESULT 7
AAR87598
ID AAR87598 standard; protein; 619 AA.
XX
XX AAR87598;
AC
XX 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 04-JUL-1996 (first entry)
XX
XX Pneumococcal surface protein (PspA).

```

```

XX PspA; pneumococcal surface protein; truncated; immunoprotective;
KW soluble fragment; insertion-duplication mutagenesis.
XX
XX Streptococcus pneumoniae; strain Rx1.
OS
XX Key Location/Qualifiers
FH Peptide 1..31
FT /label= signal_peptide
FT Protein 32..619
FT /label= mature_protein
FT Region 32..319
FT /label= alpha-helical coiled-coil region
FT /note= "contains a seven-residue periodicity"
FT Region 320..401
FT /note= "proline-rich region"
FT Region 402..421
FT /note= "repeat region"
FT Region 422..441
FT /note= "repeat region"
FT Region 442..461
FT /note= "repeat region"
FT Region 462..481
FT /note= "repeat region"
FT Region 482..501
FT /note= "repeat region"
FT Region 502..521
FT /note= "repeat region"
FT Region 522..541
FT /note= "repeat region"
FT Region 542..561
FT /note= "repeat region"
FT Region 562..581
FT /note= "repeat region"
FT Region 582..619
FT /note= "hydrophobic region starts in last repeat region
is potential membrane-spanning region"
XX
XX US5476929-A.
XX 19-DEC-1995.
XX 03-JUN-1993; 93US-00072070.
XX 15-FEB-1991; 91US-00656773.
XX 12-FEB-1992; 92US-00835698.
XX (UABR-) UAB RES FOUND.
XX
XX McDaniel LS, Yother JL, Briles DE;
XX WPI; 1996-049021/05.
DR N-PSDB; AAT08979.
XX
XX New pneumococcal surface protein A fragments - comprise proline-rich
PT region and/or repeat region, used for detection of Streptococcus
PT pneumoniae.
XX
XX Claim 1; Col 15-20; 23pp; English.
PS
XX The present sequence is that of PspA (pneumococcal surface protein A)
CC encoded by AAT08979. Through the technique of insertion-duplication
CC mutagenesis of the pspA gene of the strain Rx1 of Streptococcus
CC pneumoniae with plasmids contg. cloned fragments of the pspA structural
CC gene, it has been possible to produce soluble fragments of PspA that are
CC secreted by pneumococci. The method can be used to provide an
CC immunoprotective truncated PspA protein. Primers and probes based on the
CC present sequence are claimed, and are useful for the detection of (at
CC least 32) S. pneumoniae strains. (Updated on 25-MAR-2003 to correct PF
CC field.) (Updated on 16-OCT-2003 to standardise OS field)
XX
XX Sequence 619 AA;
SQ

```



```
Query Match      93.2%; Score 453; DB 2; Length 619;
Best Local Similarity 93.9%; Pred. No. 2.4e-32;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQLKAVE 60
    |||||
Db 223 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQLKAAE 282
    |||||

QY 61 ENNVEDYSTEGLEKTIAAKKTLEKTEADLKKAVNEPE 99
    |||||
Db 283 ENNVEDYFKGLEKTIAAKKALEKTEADLKKAVNEPE 321
    |||||

RESULT 8
AAR86911
ID AAR86911 standard; protein; 619 AA.
XX
AC AAR86911;
XX
DT 16-OCT-2003 (revised)
DT 11-MAY-1996 (first entry)
XX
DE Pneumococcal surface protein A.
XX
KW Pneumococcal surface protein A; PspA; cross-protection; vaccine;
KW Streptococcus pneumoniae; probe; primer; polymerase chain reaction;
KW otitis media; meningitis; bacteraemia; pneumonia; epitope.
XX
OS Streptococcus pneumoniae; strain Rxl.
XX
FH Key Location/Qualifiers
FT Peptide 1..31
FT /label= Sig_peptide
FT Region 32..288
FT /note= "N-terminal region is highly charged and includes
FT an alpha-helix structure"
FT Region 289..619
FT /note= "C-terminal region includes a proline-rich region
FT and a repeat region"
XX
XX AU9520112-A.
XX
XX 30-NOV-1995.
XX
XX 18-MAY-1995; 95AU-00020112.
XX
XX 20-MAY-1994; 94US-00246636.
XX 07-OCT-1994; 94US-00319795.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, Yother JL, McDaniel LS;
XX
XX WPI; 1996-030801/04.
XX N-PSDB; AAT071178.
XX
XX Pneumococcal DNA primers and probes - amplify and detect cross-protective
XX epitope(s) from Streptococcus pneumoniae surface protein A.
XX
XX Disclosure; Page 41-43; 61pp; English.
XX
XX Surface protein A, PspA (AAR86911), of Streptococcus pneumoniae Rxl is
XX the product of the pspA gene (AAT07178). PspA includes regions,
XX comprising e.g. amino acids 182-588, 293-588 and 192-299, that elicit
XX cross-protection against challenge by multiple wild-type strains of S.
XX pneumoniae. These cross-reactive epitopes can be prepd. by expression of
XX DNA obtd. by PCR amplification (see AAT07179-96), for use in vaccine
XX compns. (Updated on 16-OCT-2003 to standardise OS field)
XX
XX Sequence 619 AA;

Query Match      93.2%; Score 453; DB 2; Length 619;
Best Local Similarity 93.9%; Pred. No. 2.4e-32;
```

```
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQLKAVE 60
    |||||
Db 223 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQLKAAE 282
    |||||

QY 61 ENNVEDYSTEGLEKTIAAKKTLEKTEADLKKAVNEPE 99
    |||||
Db 283 ENNVEDYFKGLEKTIAAKKALEKTEADLKKAVNEPE 321
    |||||

RESULT 9
AAY41838
ID AAY41838 standard; protein; 619 AA.
XX
AC AAY41838;
XX
DT 08-DEC-1999 (first entry)
XX
DE Streptococcus pneumoniae Rxl PspA protein sequence.
XX
KW Streptococcus pneumoniae Rxl; PspA; immunoprotective; vaccine; diagnosis;
KW infection; pneumococcal surface protein A.
XX
OS Streptococcus pneumoniae.
XX
XX US5965400-A.
XX
XX 12-OCT-1999.
XX
XX 23-MAY-1994; 94US-00247491.
XX
XX 15-FEB-1991; 91US-00656773.
XX 12-FEB-1992; 92US-00835698.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Yother JL, Briles DE;
XX WPI; 1999-579913/49.
XX N-PSDB; AA225063.
XX
XX DNA encoding a truncated pneumococcal surface protein A used in the
XX development of pneumococcal infections.
XX
XX Claim 1; Fig 3; 27pp; English.
XX
XX The present sequence represents Streptococcus pneumoniae Rxl
XX immunoprotective Pneumococcal surface protein A (PspA). The present
XX invention also describes a method of forming the immunoprotective
XX truncated PspA, comprising incorporating a vector comprising the isolated
XX DNA molecule encoding PspA (I), into a bacterium via transfection. (I)
XX is used to design primers which are capable of detecting a large number
XX of S. pneumoniae strains, which in turn can be used to diagnose
XX pneumococcal infection in mammals (e.g. humans), independent of the
XX strain which has caused it. The PspA protein is used to develop a vaccine
XX against pneumococcal infection comprising, as an immunologically-active
XX component, a live attenuated or killed bacteria containing a gene coding
XX for the truncated form of PspA
XX
XX Sequence 619 AA;

Query Match      93.2%; Score 453; DB 2; Length 619;
Best Local Similarity 93.9%; Pred. No. 2.4e-32;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQLKAVE 60
    |||||
Db 223 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQLKAAE 282
    |||||

QY 61 ENNVEDYSTEGLEKTIAAKKTLEKTEADLKKAVNEPE 99
    |||||
Db 283 ENNVEDYFKGLEKTIAAKKALEKTEADLKKAVNEPE 321
    |||||
```

```

RESULT 10
AAE18782
ID AAE18782 standard; protein; 619 AA.
XX AC AAE18782;
XX DT 17-MAY-2002 (first entry)
XX DE S. pneumoniae Rx1 strain pneumococcal surface protein A (PspA).
XX KW Coiled-coil structural scaffold; heptad repeat; epitope; immune response;
XX KW cell-mediated immunity; microbial infection; cross-protection; therapy;
XX KW antimicrobial; vaccine; pneumococcal surface protein A; PspA.
XX OS Streptococcus pneumoniae.
XX FH Key Location/Qualifiers
FT Domain 1..314
FT /label= Helical_domain
FT Region 1..303
FT /note= "N-terminal region"
FT Region 38..44
FT /note= "Immunogenic region 3"
FT Region 40..46
FT /note= "Immunogenic region 5"
FT Region 75..80
FT /note= "Immunogenic region 29"
FT Region 82..87
FT /note= "Immunogenic region 52"
FT Region 96..101
FT /note= "Immunogenic region 66"
FT Region 114..119
FT /note= "Immunogenic region 73"
FT Region 130..135
FT /note= "Immunogenic region 78"
FT Region 137..142
FT /note= "Immunogenic region 89"
FT Region 140..145
FT /note= "Immunogenic region 91"
FT Region 152..156
FT /note= "Immunogenic region 95"
FT Domain 153..170
FT /label= Coiled_coil_motif
FT Region 161..164
FT /note= "Immunogenic region 101"
FT Region 166..170
FT /note= "Immunogenic region 116"
FT Region 173..177
FT /note= "Immunogenic region 122"
FT Region 176..180
FT /note= "Immunogenic region 123"
FT Domain 181..198
FT /label= Coiled_coil_motif
FT Region 187..191
FT /note= "Immunogenic region 130"
FT Region 194..198
FT /note= "Immunogenic region 133"
FT Region 215..219
FT /note= "Immunogenic region 140"
FT Region 226..230
FT /note= "Immunogenic region 145"
FT Region 229..232
FT /note= "Immunogenic region 148"
FT Region 260..264
FT /note= "Immunogenic region 166"
FT Region 263..267
FT /note= "Immunogenic region 168"
FT Region 284..287
FT /note= "Immunogenic region 76"
FT Region 286..290
FT /note= "Immunogenic region 179"

```

```

FT Region 294..298
FT /note= "Immunogenic region 182"
FT Region 295..299
FT /note= "Immunogenic region 185"
FT Region 309..313
FT /note= "Immunogenic region 195"
FT Region 322..326
FT /note= "Immunogenic region 206"
XX WO200196368-A2.
XX PD 20-DEC-2001.
XX 14-JUN-2001; 2001WO-US019168.
XX 14-JUN-2000; 2000US-0211892P.
XX 23-JUN-2000; 2000US-0213387P.
XX (CYTO-) CYTOVAX BIOTECHNOLOGIES INC.
XX Houston ME, Hodges RS;
XX WPI; 2002-188298/24.
XX New synthetic peptide derived from naturally occurring microbial and non-
XX microbial protein antigen useful to stimulate and elicit an immune
XX response in an animal.
XX Example 1; Page 90-92; 99pp; English.
XX The invention relates to the uses of coiled-coil structural scaffold to
XX generate structure-specific peptides, including synthetic peptides
XX derived from naturally occurring microbial and non-microbial protein
XX antigens. The structure of the synthetic peptides utilizes a scaffold of
XX heptad repeat units into which epitopes derived from coiled-coil regions
XX of native proteins are spliced. The resulting peptide has a more stable
XX coiled-coil structure, hence improving presentation of the epitopes in a
XX helical conformation. The peptides of the invention are used to stimulate
XX and elicit an immune response in an animal, as vaccine, to treat or
XX prevent microbial infection by several strains and/or species of
XX microorganism, to provide cross-protection to at least one strain and/or
XX species of microorganism and to stimulate antibody production or cell-
XX mediated immunity to the naturally occurring protein. The present
XX sequence is Streptococcus pneumoniae Rx1 strain pneumococcal surface
XX protein A (PspA) which adopts a coiled-coil structure
XX SQ Sequence 619 AA;
Query Match 93.2%; Score 453; DB 5; Length 619;
Best Local Similarity 93.9%; Pred. No. 2.4e-32;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 1 LKEIDSESDYAKGFRAPLHSLKDAKQKLSKLESLSDKIDELDAETAKLEDQKAVE 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDAKAKLSKLESLSDKIDELDAETAKLEDQKAAE 282
Qy 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKKAVNEPE 99
Db 283 ENNVEDYFKGLEKTIAAKAELEKTEADLKKAVNEPE 321
RESULT 11
ABU45778
ID ABU45778 standard; protein; 619 AA.
XX AC ABU45778;
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #31305.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX

```

```
OS Streptococcus pneumoniae.
PN WO200277183-A2.
XX
XX
PD 03-OCT-2002.
XX
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-0299926/02.
DR N-PSDB; ACA49648.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 73702; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway;
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 619 AA;
SQ
Query Match 93.2%; Score 453; DB 6; Length 619;
Best Local Similarity 93.9%; Pred. No. 2.4e-32;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 LKEIDSESDYAKGFRAPLHSLDKAKQKLSKLEELSDKIDELDAETAKLEDQKAVE 60
DB 223 LKEIDSESDYAKGFRAPLHSLDKAKQKLSKLEELSDKIDELDAETAKLEDQKAVE 282
QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKKAVNEPE 99
DB 283 ENNVEDYFKGLEKTIAAKKAELKTEADLKKAVNEPE 321
RESULT 13
AAW70336
ID AAW70336 standard; protein; 648 AA.
XX
```

```
RESULT 12
ADO52126
ID ADO52126 standard; protein; 619 AA.
XX
XX ADO52126;
AC
XX 12-AUG-2004 (first entry)
DT
XX Streptococcus pneumoniae Rx1 PspA protein.
DE
XX Immunogenic composition; vaccine; Th2-type immune response;
KW pneumococcal surface protein A; PspA.
XX
XX Streptococcus pneumoniae.
OS
XX Key Location/Qualifiers
FH Peptide 1..31
FT /label= signal_peptide
FT Protein 32..619
FT /note= "S. pneumoniae Rx1 mature PspA protein"
XX
XX US2004101531-A1.
PN
XX 27-MAY-2004.
PD
XX
XX 15-APR-2003; 2003US-00414532.
XX
XX 16-APR-2002; 2002US-0372710P.
PR
XX (CURT/) CURTISS R.
PA (KANG/) KANG H Y.
XX
XX Curtiss R, Kang HY;
XX WPI; 2004-399655/37.
DR N-PSDB; ADO52125.
XX
XX New vaccine comprising a live attenuated strain of pathogenic gram-
PT negative bacteria, useful in eliciting a Th2-type immune response in a
PT vertebrate against pathogens, e.g., helminths, fungi, viruses, protozoans
PT or bacteria.
XX
XX Example 8; SEQ ID NO 72; 94pp; English.
XX
XX The invention relates to immunogenic compositions and vaccines comprising
CC a live attenuated strain of pathogenic gram negative bacteria that
CC secretes an antigen. The vaccine is useful in eliciting a Th2-type immune
CC response in a vertebrate against pathogens, e.g., helminths, fungi,
CC viruses, protozoans or bacteria. The present sequence is Streptococcus
CC pneumoniae Rx1 pneumococcal surface protein A (PspA). This sequence is
CC used in the exemplification of the invention.
XX
XX Sequence 619 AA;
SQ
Query Match 93.2%; Score 453; DB 8; Length 619;
Best Local Similarity 93.9%; Pred. No. 2.4e-32;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 LKEIDSESDYAKGFRAPLHSLDKAKQKLSKLEELSDKIDELDAETAKLEDQKAVE 60
DB 223 LKEIDSESDYAKGFRAPLHSLDKAKQKLSKLEELSDKIDELDAETAKLEDQKAVE 282
QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKKAVNEPE 99
DB 283 ENNVEDYFKGLEKTIAAKKAELKTEADLKKAVNEPE 321
RESULT 13
AAW70336
ID AAW70336 standard; protein; 648 AA.
XX
```


XX AC AAY41837;
XX DT 08-DEC-1999 (first entry)
XX DE Streptococcus pneumoniae Rxl PspA protein sequence.
XX KW Streptococcus pneumoniae Rxl; PspA; immunoprotective; vaccine; diagnosis;
XX KW infection; pneumococcal surface protein A.
XX OS Streptococcus pneumoniae.

Search completed: June 21, 2005, 10:10:14
Job time : 74.8459 secs

XX PH Key Location/Qualifiers
XX FT Misc-difference 619. .620
XX FT /note= "a stop codon is present in the nucleotide
XX FT sequence at this position"
XX FT Misc-difference 621. .622
XX FT /note= "a stop codon is present in the nucleotide
XX FT sequence at this position"
XX FT Misc-difference 625. .626
XX FT /note= "a stop codon is present in the nucleotide
XX FT sequence at this position"
XX FT Misc-difference 630. .631
XX FT /note= "a stop codon is present in the nucleotide
XX FT sequence at this position"
XX FT Misc-difference 632. .633
XX FT /note= "a stop codon is present in the nucleotide
XX FT sequence at this position"

XX PN US5965400-A.
XX PD 12-OCT-1999.
XX XX 23-MAY-1994; 94US-00247491.
XX XX 15-FEB-1991; 91US-00656773.
XX PR 12-FEB-1992; 92US-00835698.
XX XX (UABR-) UAB RES FOUND.
XX XX Yother JL, Briles DE;
XX WPI; 1999-579913/49.
XX DR N-PSDB; AA225063.
XX PT DNA encoding a truncated pneumococcal surface protein A used in the
XX PT development of pneumococcal infections.
XX PS Claim 1; Fig 3; 27pp; English.

XX CC The present sequence represents Streptococcus pneumoniae Rxl
XX CC immunoprotective Pneumococcal surface protein A (PspA). The present
XX CC invention also describes a method of forming the immunoprotective
XX CC truncated PspA, comprising incorporating a vector comprising the isolated
XX CC DNA molecule encoding PspA (I), into a bacterium via transformation. (I)
XX CC is used to design primers which are capable of detecting a large number
XX CC of S. pneumoniae strains, which in turn can be used to diagnose
XX CC pneumococcal infection in mammals (e.g. humans), independent of the
XX CC strain which has caused it. The PspA protein is used to develop a vaccine
XX CC against pneumococcal infection comprising, as an immunologically-active
XX CC component, a live attenuated or killed bacteria containing a gene coding
XX CC for the truncated form of PspA

XX SQ Sequence 648 AA;

Query Match 93.2%; Score 453; DB 2; Length 648;
Best Local Similarity 93.9%; Pred. No. 2.5e-32;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLQSKLDAKAKLSELSKDIDELDAETAKLEDQKAVE 60
DB 223 LKEIDSESDYAKGFRAPLQSKLDAKAKLSELSKDIDELDAETAKLEDQKAAE 282

inis Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:55:14 ; Search time 18.4867 Seconds
(without alignments)
399.760 Million cell updates/sec

Title: US-10-674-755-13

Perfect score: 486

Sequence: 1 LKEIDSESEDYAKGFRAP.....KKTELEKTEADLKAVNEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents:AA:*

- 1: /cgn2_6/prodata/1/iaa/5A COMB.pep.*
- 2: /cgn2_6/prodata/1/iaa/5B COMB.pep.*
- 3: /cgn2_6/prodata/1/iaa/6A COMB.pep.*
- 4: /cgn2_6/prodata/1/iaa/6B COMB.pep.*
- 5: /cgn2_6/prodata/1/iaa/PCUTUS COMB.pep.*
- 6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	486	100.0	99	4	US-09-147-875A-13
2	482	99.2	99	2	US-08-710-749-12
3	482	99.2	195	4	US-08-529-055-71
4	482	99.2	8991	4	US-08-714-741-32
5	453	93.2	99	2	US-08-710-749-11
6	453	93.2	198	4	US-08-529-055-61
7	453	93.2	619	1	US-08-465-746-2
8	453	93.2	619	1	US-08-214-164-2
9	453	93.2	619	2	US-08-467-852A-3
10	453	93.2	619	2	US-08-246-636-2
11	453	93.2	619	2	US-08-247-491A-3
12	453	93.2	619	2	US-08-319-795-2
13	453	93.2	619	2	US-08-468-985-2
14	453	93.2	619	3	US-08-312-949-2
15	453	93.2	648	1	US-08-072-070-2
16	453	93.2	648	1	US-08-469-434-2
17	453	93.2	648	1	US-08-214-222-2
18	453	93.2	648	2	US-08-467-852A-2
19	453	93.2	648	2	US-08-468-718-2
20	453	93.2	648	2	US-08-247-491A-2
21	453	93.2	648	3	US-08-446-201-3
22	453	93.2	695	1	US-08-127-499A-23
23	453	93.2	695	1	US-08-482-847-23
24	446.5	91.9	100	4	US-09-147-875A-12
25	446	91.8	99	2	US-08-710-749-10
26	446	91.8	99	4	US-09-147-875A-11
27	446	91.8	204	4	US-08-529-055-51

28	441	90.7	288	3	US-08-312-949-4	Sequence 4, Appli
29	441	90.7	288	3	US-08-446-201-4	Sequence 4, Appli
30	430.5	88.6	289	1	US-08-072-070-4	Sequence 4, Appli
31	430.5	88.6	289	1	US-08-469-434-4	Sequence 4, Appli
32	430.5	88.6	289	1	US-08-214-222-4	Sequence 4, Appli
33	430.5	88.6	289	2	US-08-467-852A-5	Sequence 5, Appli
34	430.5	88.6	289	2	US-08-468-718-4	Sequence 4, Appli
35	430.5	88.6	289	2	US-08-247-491A-5	Sequence 5, Appli
36	407	83.7	623	4	US-08-714-741-47	Sequence 17, Appl
37	399	82.1	99	2	US-08-710-749-17	Sequence 47, Appl
38	394	81.1	1231	4	US-08-714-741-41	Sequence 10, Appl
39	389.5	80.1	100	4	US-09-147-875A-10	Sequence 10, Appl
40	388	79.8	99	2	US-08-710-749-15	Sequence 15, Appl
41	385	79.2	170	4	US-08-529-055-60	Sequence 60, Appl
42	385	79.2	181	4	US-08-529-055-42	Sequence 42, Appl
43	385	79.2	864	4	US-08-714-741-40	Sequence 40, Appl
44	382	78.6	99	4	US-09-147-875A-16	Sequence 16, Appl
45	382	78.6	188	4	US-08-529-055-59	Sequence 59, Appl

ALIGNMENTS

RESULT 1
US-09-147-875A-13
; Sequence 13, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-13

Query Match : 100.0%; Score 486; DB 4; Length 99;
Best Local Similarity 100.0%; Pred. No. 1.9e-37;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	LKEIDSESEDYAKGFRAPLHSLDKLDAKQAKLSKLEELSDKIDELDAETAKLEDQLKAVE	60
Db	1	LKEIDSESEDYAKGFRAPLHSLDKLDAKQAKLSKLEELSDKIDELDAETAKLEDQLKAVE	60
QY	61	ENNVEDYSTEGLEKTIAAKTELEKTEADLKAVNEPE	99
Db	61	ENNVEDYSTEGLEKTIAAKTELEKTEADLKAVNEPE	99

RESULT 2
US-08-710-749-12
; Sequence 12, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESS: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:

```

, MEDIUM TYPE: Floppy disk
, COMPUTER: IBM PC compatible
, OPERATING SYSTEM: PC-DOS/MS-DOS
, SOFTWARE: Patent In Release #1.0, Version #1.30
, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: US/08/710,749
, FILING DATE: 20-SEP-1996
, CLASSIFICATION: 435
, ATTORNEY/AGENT INFORMATION:
, NAME: Frommer, William S.
, REGISTRATION NUMBER: 25,506
, REFERENCE/DOCKET NUMBER: 454312-2074
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: (212) 840-3333
, TELEFAX: (212) 840-0712
, INFORMATION FOR SEQ ID NO: 12:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 99 amino acids
, TYPE: amino acid
, STRANDEDNESS: n/a
, TOPOLOGY: linear
, MOLECULE TYPE: amino acid
, US-08-710-749-12

```

Query Match	99.2%;	Score 482;	DB 2;	Length 99;
Best Local Similarity	99.0%;	Pred. No. 4.3e-37;		
Matches	98;	Conservative 1;	Mismatches 0;	Indels 0; Gaps 0;

Qy	1	LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAEIAKLEPQLKAVE	60
Db	1	LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAEIAKLEPQLKAVE	60
Qy	61	ENNVDYSTEGLEKTTAAKTELEKTEADLKKAVNEPE	99
Db	61	ENNVDYSTEGLEKTTAAKTELEKTEADLKKAVNEPE	99

```

RESULT 3
US-08-529-055-71
; Sequence 71, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (212) 840-3333
;
; TELEFAX: (212) 840-0712
;
; INFORMATION FOR SEQ ID NO: 71:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 195 amino acids
;
; TYPE: amino acid
;
; STRANDEDNESS: single
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: peptide
;
; US-08-529-055-71

```

```

Query Match      99.2%; Score 482; DB 4; Length 195;
Best Local Similarity 99.0%; Pred. NO. 9.8e-37;
Matches 99; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKIDISESDYAKGFRAPLHSLKIDAKQKLSKLELSDKIDELDAETAKLEDQLKAVE 60
    |||||
Db 1 LKIDISESDYAKGFRAPLHSLKIDAKKAKLSKLELSDKIDELDAETAKLEDQLKAVE 60
    |||||

Qy 61 ENNNVDYSTEGLKTIAAKKTLEKTEADLKKAVNEPE 99
    |||||
Db 61 ENNNVDYSTEGLKTIAAKKTLEKTEADLKKAVNEPE 99
    |||||

```

RESULT 4
US-08-714-741-32
; Sequence 32, Application US/08714741
; Patent No. 6500613

GENERAL INFORMATION:
APPLICANT: Briles, David E.
APPLICANT: McDaniel, Larry S.
APPLICANT: Swiatlo, Edwin
APPLICANT: Yother, Janet
APPLICANT: Crain, Marilyn J.
APPLICANT: Hollingshead, Susan
APPLICANT: Tart, Rebecca
APPLICANT: Brooks-Walter, Alexis
TITLE OF INVENTION: PNEUMOCOCCAL GENES, POR
TITLE OF INVENTION: EXPRESSION PRODUCTS THE
TITLE OF INVENTION: PORTIONS AND PRODUCTS
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESS: Curtis, Morris & Safford, P.C.
STREET: 530 Fifth Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.

```

,
, ZIP: 10036
,
, COMPUTER READABLE FORM:
,
, MEDIUM TYPE: Floppy disk
,
, COMPUTER: IBM PC compatible
,
, OPERATING SYSTEM: PC-DOS/MS-DOS
,
, SOFTWARE: PatentIn Release #1.0, Version #1.30
,
, CURRENT APPLICATION DATA:
,
, APPLICATION NUMBER: US/08/714,741
,
, FILING DATE: 16-SEP-1996
,
, CLASSIFICATION: 435
,
, ATTORNEY/AGENT INFORMATION:
,
, NAME: Frommer Esq., William S.
,
, REGISTRATION NUMBER: 25,506
,
, REFERENCE/DOCKET NUMBER: 454312-2460
,
, TELECOMMUNICATION INFORMATION:
,
, TELEPHONE: (212) 840-3333
,
, TELEFAX: (212) 840-0712
,
, INFORMATION FOR SEQ ID NO: 32:
,
, SEQUENCE CHARACTERISTICS:
,
, LENGTH: 8991 amino acids
,
, TYPE: amino acid
,
, STRANDEDNESS: single
,
, TOPOLOGY: linear
,
, MOLECULE TYPE: amino acid
,
, US-08-714-741-32

```

US-08-714-741-322


```
Query Match          99.2%; Score 482; DB 4; Length 8991;
Best Local Similarity 99.0%; Pred. No. 9.8e-35;
Matches 98; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSDKIDELDAETAKLEDQLKAVE 60
    |||||
Db 8797 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSDKIDELDAETAKLEDQLKAVE 8856
    |||||
QY 61 ENNVEDYSTEGLEKTIAAKTEADLKKAVNEPE 99
    |||||
Db 8857 ENNVEDYSTEGLEKTIAAKTEADLKKAVNEPE 8895

RESULT 5
US-08-710-749-11
; Sequence 11, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-710-749-11

Query Match          93.2%; Score 453; DB 2; Length 99;
Best Local Similarity 93.9%; Pred. No. 1.9e-34;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSDKIDELDAETAKLEDQLKAVE 60
    |||||
Db 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSDKIDELDAETAKLEDQLKAAE 60
    |||||
QY 61 ENNVEDYSTEGLEKTIAAKTEADLKKAVNEPE 99
    |||||
Db 61 ENNVEDYFKEGLEKTIAAKASLEKTEADLKKAVNEPE 99

RESULT 6
US-08-529-055-61
; Sequence 61, Application US/08529055
; Patent No. 6592876

Query Match          93.2%; Score 453; DB 2; Length 99;
Best Local Similarity 93.9%; Pred. No. 1.9e-34;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSDKIDELDAETAKLEDQLKAVE 60
    |||||
Db 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSDKIDELDAETAKLEDQLKAAE 60
    |||||
QY 61 ENNVEDYSTEGLEKTIAAKTEADLKKAVNEPE 99
    |||||
Db 61 ENNVEDYFKEGLEKTIAAKASLEKTEADLKKAVNEPE 99

RESULT 7
US-08-465-746-2
; Sequence 2, Application US/08465746
; Patent No. 5679768
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Yother, Janet L.
; APPLICANT: Yother, Janet S.
; APPLICANT: McDaniel, Larry S.
; TITLE OF INVENTION: EPITOPIC REGIONS OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEIN A
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
```

```
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-61

Query Match          93.2%; Score 453; DB 4; Length 198;
Best Local Similarity 93.9%; Pred. No. 4.4e-34;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSDKIDELDAETAKLEDQLKAVE 60
    |||||
Db 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSDKIDELDAETAKLEDQLKAAE 60
    |||||
QY 61 ENNVEDYSTEGLEKTIAAKTEADLKKAVNEPE 99
    |||||
Db 61 ENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 99

RESULT 7
US-08-465-746-2
; Sequence 2, Application US/08465746
; Patent No. 5679768
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Yother, Janet L.
; APPLICANT: Yother, Janet S.
; APPLICANT: McDaniel, Larry S.
; TITLE OF INVENTION: EPITOPIC REGIONS OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEIN A
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
```

Query Match	93.2%	Score 453;	DB 2;	Length 619;
Best Local Similarity	93.9%;	Pred. No. 1.7e-33;		
Matches 93;	Conservative	1;	Mismatches 5;	Indels 0;
			Gaps	0

Qy	1	LKEIDSESESDYAKGEGFRAPLHSLKLDADQAKLSKLSELSDKIDELDAEIAKLSEDLQKAYE	60
Db	223	LKEIDSESESDYAKGEGFRAPLQSKLDADQAKLSKLSELSDKIDELDAEIAKLSEDLQKAAE	282
Qy	61	ENNNVDYDTEGLEKTTIAAKTEKTEKTRADLKKAVNPE	99
Db	283	ENNNVDYFKBGLKTTIAAKKAEKTEKTRADLKKAVNPE	321

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

Query Match      93.2%; Score 453; DB 2; Length 619;
Best Local Similarity 93.9%; Pred. No. 1.7e-33;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEDISESDEYAKGFRAPLHSLDKAQAKLSKLELSDKIDELDAEIAKLEDPOLKAVE 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 223 LKEDISESDEYAKGFRAPLQSLDKAKAKLSKLELSDKIDELDAEIAKLEDPOLKAAE 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 61 ENNVVDYSTGLEGSKTTAAKTELEKTEADLKKAVNEPE 99
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 283 ENNVVDYFKEGLESKTTAAKKAELKEKTEADLKKAVNEPE 321
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 11

```

US-08-247-491A-3
; Sequence 3, Application US/08247491A
; Patent No. 5965400
; GENERAL INFORMATION:
; APPLICANT: BRILES, David E.
; APPLICANT: YOTHER, Janet L.
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP
; STREET: 745 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10151
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/247,491A
; FILING DATE: 23-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOWALSKI, Thomas J.
; REGISTRATION NUMBER: 32,147
; REFERENCE/DOCKET NUMBER: 454312-2041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-588-0800
; TELEFAX: 212-588-0500
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-247-491A-3

Query Match          93.2%; Score 453; DB 2; Length 619;
Best Local Similarity 93.3%; Pred. No. 1.7e-33;
Matches 93; Conservative 1; Mismatches 5; Indels

Qy      1 LKEDISESDYAKEGFRAPLHSLDKAKLSKLBEISDKIDELDAEIAK
        |||||
Db      223 LKEDISESDYAKEGFRAPLQSKLDKAKLSKLBEISDKIDELDAEIAK
        |||||

Qy      61 ENNVEDYPEGLEKTIAAKKTEADLKAVNEPE 99
        |||||
Db      283 ENNVEDYPKEGLEKTIAAKKAELETKADLKAVNEPE 321
        |||||

RESULT 12
US-08-319-795-2
; Sequence 2, Application US/08319795
; Patent No. 5980909
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Yother, Janet L.
; APPLICANT: McDaniel, Larry S
; TITLE OF INVENTION: Epitopic Regions of Pneumococcal Surface
; TITLE OF INVENTION: Protein A
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheomaker and Mattare, Ltd.
; STREET: 1203 Crystal Plaza Bldg. 1, 2001 Jefferson
; STREET: Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/319,795
; FILING DATE:
; CLASSIFICATION: 424
; APPLICATION NUMBER: US 08/246,636
; FILING DATE: 20-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/048,896
; FILING DATE: 20-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,698
; FILING DATE: 12-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/656,773
; FILING DATE: 15-FEB-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-319-795-2

Query Match          93.2%; Score 453; DB 2; Length 619;
Best Local Similarity 93.9%; Pred. No. 1.7e-33;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAKQKLSKLELSKIDELDAEIAKLEDLKAVE 60
Db 223 LKEIDSESDYAKGFRAPLHSLDKAKQKLSKLELSKIDELDAEIAKLEDLKAVE 282

Qy 61 ENNVEDYSTGEGLEKTTAAKTEADLKKA VNEPE 99
Db 283 ENNVEDYFKEGLEKTTAAKKALEKTEADLKKA VNEPE 321

RESULT 13
US-08-468-985-2
; Sequence 2, Application US/08468985
; Patent No. 5997882
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Yother, Janet L.
; APPLICANT: McDaniel, Larry S
; TITLE OF INVENTION: Epitopic Regions of Pneumococcal Surface
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheomaker and Mattare, Ltd.
; STREET: 1203 Crystal Plaza Bldg. 1, 2001 Jefferson
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,985
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: 08/319,795
; FILING DATE:
; APPLICATION NUMBER: US 08/246,636
; FILING DATE: 20-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/048,896
; FILING DATE: 20-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,698
; FILING DATE: 12-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/656,773
; FILING DATE: 15-FEB-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-985-2

Query Match          93.2%; Score 453; DB 2; Length 619;
Best Local Similarity 93.9%; Pred. No. 1.7e-33;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAKQKLSKLELSKIDELDAEIAKLEDLKAVE 60
Db 223 LKEIDSESDYAKGFRAPLHSLDKAKQKLSKLELSKIDELDAEIAKLEDLKAVE 282

Qy 61 ENNVEDYSTGEGLEKTTAAKTEADLKKA VNEPE 99
Db 283 ENNVEDYFKEGLEKTTAAKKALEKTEADLKKA VNEPE 321

RESULT 14
US-08-312-949-2
; Sequence 2, Application US/08312949
; Patent No. 6027734
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Wu, Hong-Yin
; TITLE OF INVENTION: MUCOSAL ADMINISTRATION OF
; TITLE OF INVENTION: PNEUMOCOCCAL ANTIGENS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,949
; FILING DATE: 30-SEP-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2049
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids

```

Qy 61 ENNVEDYSTEGLEKTIAAKKTELEKTEADLKAVNEPE 99

This Page Blank (uspto)

GenCore version 5.1.6
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 10:12:15 ; Search time 63.2388 Seconds
 (without alignments)
 601.118 Million cell updates/sec

Title: US-10-674-755-13

Perfect score: 486

Sequence: 1 LKEIDSESDYAKGFRAP.....KXTELEKTEADLKAVNEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	486	100.0	99	15	US-10-674-755-13
2	482	99.2	195	15	US-10-299-636-86
3	453	93.2	138	15	US-10-299-636-76
4	453	93.2	354	15	US-10-299-636-105
5	453	93.2	588	15	US-10-299-636-96
6	453	93.2	619	10	US-09-882-774-1
7	453	93.2	619	10	US-10-282-122A-73702
8	453	93.2	619	16	US-10-414-532-72
9	446.5	91.9	100	15	US-10-674-755-12
10	446	91.8	99	15	US-10-674-755-11
11	446	91.8	204	15	US-10-299-636-66

12	389.5	80.1	100	15	US-10-674-755-10	Sequence 10, Appl
13	385	79.2	170	15	US-10-299-636-75	Sequence 75, Appl
14	385	79.2	181	15	US-10-299-636-57	Sequence 57, Appl
15	385	79.2	643	15	US-10-299-636-95	Sequence 95, Appl
16	385	79.2	670	9	US-09-748-875-63	Sequence 63, Appl
17	385	79.2	670	10	US-09-298-523B-63	Sequence 63, Appl
18	385	79.2	690	9	US-09-748-875-61	Sequence 61, Appl
19	385	79.2	690	10	US-09-298-523B-61	Sequence 61, Appl
20	385	79.2	691	9	US-09-748-875-1	Sequence 1, Appl
21	385	79.2	691	10	US-09-298-523B-1	Sequence 1, Appl
22	385	79.2	701	9	US-09-748-875-62	Sequence 62, Appl
23	385	79.2	701	10	US-09-298-523B-62	Sequence 62, Appl
24	385	79.2	707	9	US-09-748-875-2	Sequence 2, Appl
25	385	79.2	707	10	US-09-298-523B-2	Sequence 2, Appl
26	385	79.2	711	9	US-09-748-875-3	Sequence 3, Appl
27	385	79.2	711	10	US-09-298-523B-3	Sequence 3, Appl
28	385	79.2	739	17	US-10-732-923-3294	Sequence 3294, Ap
29	385	79.2	929	9	US-09-748-875-60	Sequence 60, Appl
30	385	79.2	929	10	US-09-298-523B-60	Sequence 60, Appl
31	385	79.2	929	15	US-10-299-636-94	Sequence 94, Appl
32	382	78.6	99	15	US-10-674-755-16	Sequence 16, Appl
33	382	78.6	188	15	US-10-299-636-74	Sequence 74, Appl
34	375	77.2	204	15	US-10-299-636-73	Sequence 73, Appl
35	373	76.7	141	14	US-10-254-995-2	Sequence 2, Appl
36	373	76.7	589	9	US-09-748-875-14	Sequence 14, Appl
37	373	76.7	589	10	US-09-298-523B-14	Sequence 14, Appl
38	373	76.7	589	15	US-10-299-636-97	Sequence 97, Appl
39	370	76.1	99	15	US-10-674-755-15	Sequence 15, Appl
40	358	73.7	206	15	US-10-299-636-69	Sequence 69, Appl
41	355	73.0	99	15	US-10-674-755-14	Sequence 14, Appl
42	327.5	67.4	100	15	US-10-674-755-2	Sequence 2, Appl
43	324.5	66.8	100	15	US-10-674-755-3	Sequence 3, Appl
44	322.5	66.4	194	15	US-10-299-636-79	Sequence 79, Appl
45	322	66.3	73	9	US-09-027-956-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1

US-10-674-755-13
 ; Sequence 13, Application US/10674755
 ; Publication No. US20040067237A1
 ; GENERAL INFORMATION:
 ; APPLICANT: BECKER et al.
 ; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
 ; FILE REFERENCE: 454312-2471
 ; CURRENT APPLICATION NUMBER: US/10/674,755
 ; CURRENT FILING DATE: 2003-09-30
 ; PRIOR APPLICATION NUMBER: US/09/147,875A
 ; PRIOR FILING DATE: 1999-05-24
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 13
 ; LENGTH: 99
 ; TYPE: PRT
 ; ORGANISM: Streptococcus pneumoniae
 ; US-10-674-755-13

Query Match 100.0%; Score 486; DB 15; Length 99;
 Best Local Similarity 100.0%; Pred. No. 6.9e-32;
 Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	LKEIDSESDYAKGFRAPLHSLDKAKLSKLEELSDKIDELDAETAKLEDQKAVE	60
Db	1	LKEIDSESDYAKGFRAPLHSLDKAKLSKLEELSDKIDELDAETAKLEDQKAVE	60

QY	61	ENNVNEDYSTEGLKTIAAKTELEKTEADLKAVNEPE	99
Db	61	ENNVNEDYSTEGLKTIAAKTELEKTEADLKAVNEPE	99

RESULT 2

```
US-10-299-636-86
; Sequence 86, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-86

Query Match          99.2%; Score 482; DB 15; Length 195;
Best Local Similarity 99.0%; Pred. No. 3.1e-31;
Matches 98; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSDKIDELDAEIAKLEDLQKAVE 60
Db 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSDKIDELDAEIAKLEDLQKAVE 60

Qy 61 ENNVEDYSTEGLEKTIAAKKTEADLKKA VNEPE 99
Db 61 ENNVEDYSTEGLEKTIAAKKTEADLKKA VNEPE 99

US-10-299-636-76
; Sequence 76, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-76

Query Match          93.2%; Score 453; DB 15; Length 198;
Best Local Similarity 93.9%; Pred. No. 6.8e-29;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
Qy 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSDKIDELDAEIAKLEDLQKAVE 60
Db 1 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLELSDKIDELDAEIAKLEDLQKAAE 60

Qy 61 ENNVEDYSTEGLEKTIAAKKTEADLKKA VNEPE 99
Db 61 ENNVEDYFKEGLEKTIAAKKA E LKTEADLKKA VNEPE 99

RESULT 4
US-10-299-636-105
; Sequence 105, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 105
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-105

Query Match          93.2%; Score 453; DB 15; Length 354;
Best Local Similarity 93.9%; Pred. No. 1.3e-28;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSDKIDELDAEIAKLEDLQKAVE 60
Db 175 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLELSDKIDELDAEIAKLEDLQKAAE 234

Qy 61 ENNVEDYSTEGLEKTIAAKKTEADLKKA VNEPE 99
Db 235 ENNVEDYFKEGLEKTIAAKKA E LKTEADLKKA VNEPE 273

RESULT 5
US-10-299-636-96
; Sequence 96, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
```



```
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 96
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-96

Query Match          93.2%; Score 453; DB 15; Length 588;
Best Local Similarity 93.9%; Pred. No. 2.3e-28;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
Db 192 LKEIDSESDYAKGFRAPLQSKLDKAKKAKLSKLELSKIDELDAEIAKLEDLQKAAE 251

QY 61 ENNVEDYSTGEGLEKTIAAKTELEKTEADLKKAVNEPE 99
Db 252 ENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 290

RESULT 6
US-09-882-774-1
; Sequence 1, Application US/09882774
; Publication No. US20030021795A1
; GENERAL INFORMATION:
; APPLICANT: Houston, Michael E.
; APPLICANT: Hodges, Robert
; TITLE OF INVENTION: Use of Coiled-Coil Structural Scaffold to Generate
; TITLE OF INVENTION: Structure-Specific Peptides
; FILE REFERENCE: 003592-007
; CURRENT APPLICATION NUMBER: US/09/882,774
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,892
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/213,387
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-882-774-1

Query Match          93.2%; Score 453; DB 10; Length 619;
Best Local Similarity 93.9%; Pred. No. 2.4e-28;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDKAKKAKLSKLELSKIDELDAEIAKLEDLQKAAE 282

QY 61 ENNVEDYSTGEGLEKTIAAKTELEKTEADLKKAVNEPE 99
Db 283 ENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 321

RESULT 7
US-10-282-122A-73702
; Sequence 73702, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert

Query Match          93.2%; Score 453; DB 16; Length 619;
Best Local Similarity 93.9%; Pred. No. 2.4e-28;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDKAKKAKLSKLELSKIDELDAEIAKLEDLQKAAE 282

QY 61 ENNVEDYSTGEGLEKTIAAKTELEKTEADLKKAVNEPE 99
Db 283 ENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 321

RESULT 8
US-10-414-532-72
; Sequence 72, Application US/10414532
; Publication No. US20040101531A1
; GENERAL INFORMATION:
; APPLICANT: CURTISS III, ROY
; APPLICANT: KANG, HO YOUNG
; TITLE OF INVENTION: IMPROVED IMMUNOGENIC COMPOSITIONS AND VACCINES COMPRISING
; FILE REFERENCE: 56029-40437
; CURRENT APPLICATION NUMBER: US/10/414,532
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: 60/372,710
; PRIOR FILING DATE: 2002-04-16
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 72
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-414-532-72

Query Match          93.2%; Score 453; DB 16; Length 619;
Best Local Similarity 93.9%; Pred. No. 2.4e-28;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAQAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDAKAKLSKLELSKIDELDAEIAKLEDLQKAAE 282

Qy 61 ENNVEDYSTEGLEKTTAAKTELEKTEADLKKAVNEPE 99
Db 283 ENNVEDYFKGLEKTTAAKAELEKTEADLKKAVNEPE 321

RESULT 9
US-10-674-755-12
; Sequence 12, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 100
; TYPE: PRP
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-12

Query Match 91.9%; Score 446.5; DB 15; Length 100;
Best Local Similarity 94.0%; Pred. No. 1.1e-28;
Matches 94; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAQAKLSKLELSKIDELDAEIAKLEDLQKAV 59
Db 1 LKEIDSESDYAKGFRAPLQSKLDAKAKLSKLELSKIDELDAEIAKLEDLQKAA 60

Qy 60 ENNVEDYSTEGLEKTTAAKTELEKTEADLKKAVNEPE 99
Db 61 ENNVEDYFKGLEKTTAAKAELEKTEADLKKAVNEPE 100

RESULT 10
US-10-674-755-11
; Sequence 11, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 99
; TYPE: PRP
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-11

Query Match 91.8%; Score 446; DB 15; Length 99;
Best Local Similarity 91.9%; Pred. No. 1.1e-28;
Matches 91; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAQAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
Db 1 LKEIDSESDYKVEGFRAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEDLQKAAE 60

Qy 61 ENNVEDYSTEGLEKTTAAKTELEKTEADLKKAVNEPE 99
Db 61 ENNVEDYFKGLEKTTAAKAELEKTEADLKKAVNEPE 99

RESULT 11
US-10-674-755-10
; Sequence 10, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Iocher, Janec
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 66
; LENGTH: 204
; TYPE: PRP
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-66

Query Match 91.8%; Score 446; DB 15; Length 204;
Best Local Similarity 91.9%; Pred. No. 2.6e-28;
Matches 91; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAQAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
Db 1 LKEIDSESDYKVEGFRAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEDLQKAAE 60

Qy 61 ENNVEDYSTEGLEKTTAAKTELEKTEADLKKAVNEPE 99
Db 61 ENNVEDYFKGLEKTTAAKAELEKTEADLKKAVNEPE 99

RESULT 12
US-10-674-755-10
; Sequence 10, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 100
; TYPE: PRP
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-10

Query Match 80.1%; Score 389.5; DB 15; Length 100;
Best Local Similarity 84.0%; Pred. No. 4.1e-24;
Matches 84; Conservative 3; Mismatches 12; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAQAKLSKLELSKIDELDAEIAKLEDLQKAV 59
Db 1 LKEIDSESDYKVEGFRAPLQSKLDAKAKLSKLELSKIDELDAEIAKLEDLQKDA 60

Db 61 ENNVEDYFKGLEKTTAAKAELEKTEADLKKAVNEPE 99

RESULT 11
US-10-299-636-66
; Sequence 66, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Iocher, Janec
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 66
; LENGTH: 204
; TYPE: PRP
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-66

Query Match 91.8%; Score 446; DB 15; Length 204;
Best Local Similarity 91.9%; Pred. No. 2.6e-28;
Matches 91; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAQAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
Db 1 LKEIDSESDYKVEGFRAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEDLQKAAE 60

Qy 61 ENNVEDYSTEGLEKTTAAKTELEKTEADLKKAVNEPE 99
Db 61 ENNVEDYFKGLEKTTAAKAELEKTEADLKKAVNEPE 99

RESULT 12
US-10-674-755-10
; Sequence 10, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 100
; TYPE: PRP
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-10

Query Match 80.1%; Score 389.5; DB 15; Length 100;
Best Local Similarity 84.0%; Pred. No. 4.1e-24;
Matches 84; Conservative 3; Mismatches 12; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAQAKLSKLELSKIDELDAEIAKLEDLQKAV 59
Db 1 LKEIDSESDYKVEGFRAPLQSKLDAKAKLSKLELSKIDELDAEIAKLEDLQKDA 60

QY 60 ENNVEDYSTEGLEKTIAAKTELEKTEADLKAVNEPE 99
Db 61 EGNNVEAYFKEGLEKTTAEKKAELKAEADLKAVDEPE 100

RESULT 13

US-10-299-636-75
; Sequence 75, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299, 636
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-75

Query Match 79.2%; Score 385; DB 15; Length 170;
Best Local Similarity 81.8%; Pred. No. 1.7e-23;
Matches 81; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLEELSDKIDELDAEIAKLEDLKAVE 60
Db 1 LKEIDSDSDYKGLRAPLQSKLDTKKAQKLSKLEELSDKIDELDAEIAKLEVLKDAE 60

QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKAVNEPE 99
Db 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKAVDEPE 99

RESULT 14

US-10-299-636-57
; Sequence 57, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299, 636
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
; LENGTH: 181
; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae
US-10-299-636-57

Query Match 79.2%; Score 385; DB 15; Length 181;
Best Local Similarity 81.8%; Pred. No. 1.8e-23;
Matches 81; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLEELSDKIDELDAEIAKLEDLKAVE 60
Db 1 LKEIDSDSDYKGLRAPLQSKLDTKKAQKLSKLEELSDKIDELDAEIAKLEVLKDAE 60
QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKAVNEPE 99
Db 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKAVDEPE 99

RESULT 15

US-10-299-636-95
; Sequence 95, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299, 636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 95
; LENGTH: 643
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-95

Query Match 79.2%; Score 385; DB 15; Length 643;
Best Local Similarity 81.8%; Pred. No. 7.5e-23;
Matches 81; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLEELSDKIDELDAEIAKLEDLKAVE 60
Db 245 LKEIDSDSDYKGLRAPLQSKLDTKKAQKLSKLEELSDKIDELDAEIAKLEVLKDAE 304

QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKAVNEPE 99
Db 305 GNNVAYFKEGLEKTTAEKKAELKAEADLKAVDEPE 343

Search completed: June 21, 2005, 11:18:34
Job time : 63.2388 sec

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:53:39 ; Search time 9.9 Seconds
(without alignments)
962.168 Million cell updates/sec

Title: US-10-674-755-13

Perfect score: 486

Sequence: 1 LKEIDESESDYAKGFRAP.....KKTELEKTEADLKKA VNEPE 99

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79: *
1: Pirl: *
2: Pirl: *
3: Pirl: *
4: Pirl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	453	93.2	619	2 A97887	surface protein ps
2	453	93.2	619	2 A41971	surface protein ps
3	129.5	26.6	744	2 F95013	pneumococcal surfa
4	111.5	22.9	3488	2 T34418	hypothetical prote
5	110	22.6	281	2 F75216	hypothetical prote
6	109.5	22.5	852	2 D72230	conserved hypothet
7	108	22.2	886	2 H69378	conserved hypothet
8	107.5	22.1	161	2 S48396	tropomyosin TPM2 -
9	106.5	21.9	388	2 S52536	fcrA 15 protein -
10	106.5	21.9	405	2 A33339	Fc gamma (15G) rec
11	105	21.6	229	2 S70532	outer surface prot
12	105	21.6	764	2 T05049	hypothetical prote
13	104.5	21.5	372	2 S23326	gene ML2.2 protein
14	104.5	21.5	415	2 S35760	fcrA protein precu
15	104.5	21.5	501	2 A4643	M protein precu
16	103.5	21.3	387	2 S57834	fcrA protein precu
17	102.5	21.1	388	2 A46173	Mrp4 protein - Str
18	101	20.8	402	2 S37046	IgA receptor - Str
19	100.5	20.7	1312	2 T30845	probable DNA repai
20	100	20.6	166	2 S73342	hypothetical prote
21	100	20.6	365	2 B54128	Fc-binding protein
22	100	20.6	376	2 A43528	protein H precu
23	100	20.6	1957	2 T38077	hypothetical coile
24	99.5	20.5	1277	2 S53043	probable membrane
25	98.5	20.3	583	1 A46127	radixin - human
26	98.5	20.3	583	1 S39805	radixin - pig
27	98.5	20.3	1790	2 S67593	transport protein
28	98.5	20.3	2401	2 T28676	thropy protein -
29	98	20.2	473	2 F70031	cell wall-binding

30	98	20.2	1875	2 S38173	myosin-like protei
31	98	20.2	1937	2 I38055	myosin heavy chain
32	98	20.2	1957	2 A59294	skeletal myosin -
33	98	20.2	1979	2 C71622	hypothetical prote
34	97.5	20.1	384	2 S49550	M-like protein emm
35	97.5	20.1	880	2 F75103	conserved hypothet
36	97.5	20.1	896	2 S43074	epidermal growth f
37	97	20.0	1039	2 S62509	probable vesicular
38	97	20.0	1044	2 T50213	probable vesicular
39	97	20.0	1156	2 B70356	chromosome assembl
40	97	20.0	1177	2 B75150	glucose segrega
41	97	20.0	1319	2 A28313	glued protein - fr
42	97	20.0	1940	1 S04090	myosin heavy chain
43	96.5	19.9	230	2 I40287	outer surface prot
44	96.5	19.9	408	2 S30283	protein M precu
45	96.5	19.9	713	2 B84583	hypothetical prote

ALIGNMENTS

RESULT 1

A97887
surface protein pspA precursor [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: A97887
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; Mc
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A>Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: A97887
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-619 <KUR>
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:AE007317; PIDN:AAK98925.1; PID:gl
C:Genetics:
A:Gene: pspA

Query Match 93.2%; Score 453; DB 2; Length 619;
Best Local Similarity 93.9%; Pred. No. 9.1e-24;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY	1	LKEIDESESDYAKGFRAPLHSDAKAKLSKLEELSDKIDELDAEIAKLEDLQKAVE	60
DB	223	LKEIDESESDYAKGFRAPLQSKLDKAKAKLSKLEELSDKIDELDAEIAKLEDLQKAAE	282
QY	61	ENNVEDYSTEGLKTTAAKTELEKTEADLKKA VNEPE	99
DB	283	ENNVEDYFKEGLKTTAAKAELEKTEADLKKA VNEPE	321

RESULT 2

A41971
surface protein pspA precursor - Streptococcus pneumoniae
N:Alternate names: pneumococcal surface protein A
C:Species: Streptococcus pneumoniae
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A41971; A60282; A33134
R:Yother, J.; Briles, D.E.
J. Bacteriol. 174, 601-609, 1992
A>Title: Structural properties and evolutionary relationships of PspA, a surface protein

A:Reference number: A41971; MUID:92105030; PMID:1729249
A:Accession: A41971
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-619 <YOT>
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:M74122; NID:q153840; PIDN:AAA2701f
A>Note: sequence extracted from NCBI backbone (NCIN:75635, NCBI:P:75636)
R:Talkington, D.F.; Crammins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.

Infect. Immun. 59, 1285-1289, 1991
A;Title: A 43-kilodalton pneumococcal surface protein, PspA: isolation, protective ability
A;Reference number: A60282; MUID:91169598; PMID:2004810
A;Accession: A60282
A;Molecule type: protein
A;Residues: 32-76 <TAL>
A;Experimental source: strain JY2008
C;Genetics:
A;Gene: pspA
F;1-31/Domain: signal sequence #status predicted <SIG>
F;32-619/Product: surface protein pspA #status predicted <MAT>
F;411-430/Domain: cpl repeat homology <CP01>
F;431-450/Domain: cpl repeat homology <CP02>
F;451-470/Domain: cpl repeat homology <CP03>
F;471-490/Domain: cpl repeat homology <CP04>
F;491-510/Domain: cpl repeat homology <CP05>
F;511-530/Domain: cpl repeat homology <CP06>
F;531-550/Domain: cpl repeat homology <CP07>
F;551-570/Domain: cpl repeat homology <CP08>
F;571-591/Domain: cpl repeat homology <CP09>
F;592-611/Domain: cpl repeat homology <CP10>

Query Match 93.2%; Score 453; DB 2; Length 619;
Best Local Similarity 93.9%; Pred. No. 9.1e-24;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDSESEDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAEIAKLELDQKAVE 60
Db 223 LKEIDSESEDYAKGFRAPLQSLDKDAKQAKLSKLELSDKIDELDAEIAKLELDQKAAE 282

Qy 61 ENNVEDYSTGLEKTIAAKTELEKTEADLKKAVNEPE 99
Db 283 ENNVEDYFKEGLSEKTIAAKKAELKTEADLKKAVNEPE 321

RESULT 3
F95013
pneumococcal surface protein A [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: F95013
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
son, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,
n, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: F95013
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-744 <KUR>
A;Cross-references: UNIPROT:Q97T39; GB:AB005672; PIDN:AAK74303.1; PID:gl4971584; GSPDB:G
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP0117

Query Match 26.6%; Score 129.5; DB 2; Length 744;
Best Local Similarity 36.1%; Pred. No. 0.11;
Matches 39; Conservative 12; Mismatches 28; Indels 29; Gaps 3;

Qy 19 APLHSLDKDAKQAKLSK-----LEELSDKI-----DELDAEIAKLELDQKAVEENNVED 67
Db 333 AALQNLAAKKAELAKQTELEKLLSDLPBGKTQDELKREAEAEALDKKADLQNKVAD 392

Qy 68 Y-----STEGLEKTIAAKTELEKTEADLKKAVNE 97
Db 393 LEKEISNLEILLGADSEDDTAALQNLKATKKAELEKTKQKELDAALNE 440

RESULT 4
T34418
hypothetical protein F12F3.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T34418
R;Fulton, B.; Wohldmann, P.
submitted to the EMBL Data Library, July 1998
A;Description: The sequence of C. elegans cosmid F12F3.
A;Reference number: Z21521
A;Accession: T34418
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-3488 <FUL>
A;Cross-references: EMBL:U80022; PIDN:AAC25885.1; GSPDB:GN00023; CESP:F12F3.3
A;Experimental source: strain Bristol N2; clone F12F3
C;Genetics:
A;Gene: CESP:F12F3.3
A;Map position: 5
A;Introns: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1

Query Match 22.9%; Score 111.5; DB 2; Length 3488;
Best Local Similarity 35.2%; Pred. No. 8.4;
Matches 44; Conservative 17; Mismatches 29; Indels 35; Gaps 7;

Qy 2 KEIDES-----ESEDYAKGFRAPLHSLDKDAKQAKLSK-----EELSDKIDELDAE 49
Db 1009 KETDEKILKLDAAEIAAKTKQEADEKSKLDA-QEKIKKVSDDAARKEKELNDKL-KLSEI 1066

Qy 50 A-----KLQDLKA-----VEENNVEDYSTEGLEKTIAAKTELEKTEA 89
Db 1067 ATKASADKLKEBQQAQKAAVEAAKQKQKEDQLKDLTEAAKAAAEKLELEK-QA 1125

Qy 90 DLKKA 94
Db 1126 QIKKA 1130

RESULT 5
F75216
hypothetical protein PAR2181 - Pyrococcus abyssi (strain Orsay)
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: F75216
R;anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A;Reference number: A75001
A;Accession: F75216
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-281 <KAW>
A;Cross-references: UNIPROT:Q9V217; GB:AJ248283; GB:AL096836; NID:95457433; PIDN:CAB49181
A;Experimental source: strain Orsay
C;Genetics:
A;Gene: PAR2181

Query Match 22.6%; Score 110; DB 2; Length 281;
Best Local Similarity 27.9%; Pred. No. 0.85;
Matches 29; Conservative 26; Mismatches 39; Indels 10; Gaps 2;

Qy 2 KEIDSESEDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAEIAKLELDQKAVEE 61
Db 163 KEIEELKGKVEKLEQEKKELEKLEKSEVKLMYEAKAKGAEELEAKLREYEEKVREE 222

Qy 62 -----NNNVEDYSTEGLEKTIAAKTELEKTEADLKKAVNE 97
Db 223 LERKVELERSINEYETK--VKSLEKKCELEKNKVELEEVNK 264

RESULT 6
D72230
conserved hypothetical protein - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: D72230

R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.

Nature 399, 323-329, 1999

A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing of *Thermoplasma acidophilum*

A;Reference number: A72200; MUID:99287316; PMID:10360571

A;Accession: D72230

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-852 <ARN>

A;Cross-references: UNIPROT:Q9X1X1; GB:AE001806; GB:AE000512; NID:g4982196; PIDN:AAD3670

A;Experimental source: strain MSBB

C;Genetics:

A;Gene: Tm636

C;Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 22.5%; Score 109.5; DB 2; Length 852;

Best Local Similarity 29.2%; Pred. No. 2.8;

Matches 26; Conservative 22; Mismatches 30; Indels 11; Gaps 2;

QY 6 ESESDYAKGFRAPLHSLDKAKLKLKLE-----LSDKIDELDAEIAKLEDLKXAVEE 61

DB 506 EKTIEELHRLGYSEDLKLEKRLKRIEERHSISQKITAADVQISQIENQLKEIKG 565

QY 62 NNVEDYSTEGLEKTTAAKTELEKTEAD 90

DB 566 -----EIEAKRETLKEOREMDOLKSD 587

RESULT 7

H69378

conserved hypothetical protein AF1032 - Archaeoglobus fulgidus

C;Species: Archaeoglobus fulgidus

C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C;Accession: H69378

R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, R.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L. Nature 390, 364-370, 1997

A;Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S. Smith, H.O.; Woese, C.R.; Venter, J.C.

A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon *Pyrococcus furiosus*

A;Reference number: A69250; MUID:98049343; PMID:9389475

A;Accession: H69378

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-886 <KLE>

A;Cross-references: UNIPROT:O29230; GB:AE001032; GB:AE000782; NID:g2689355; PIDN:AAB9021

C;Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 22.2%; Score 108; DB 2; Length 886;

Best Local Similarity 28.9%; Pred. No. 3.7;

Matches 39; Conservative 22; Mismatches 28; Indels 46; Gaps 6;

QY 1 LKSIDSESDYAKGFRAPLHSLDKAKLKLKLELSDKIDELDAE----- 49

DB 303 LRDVKEKREG-DLTREA--AGIQALQKAEEDNSKLEITKRIEELERLEFEKSHRLLE 359

QY 50 -----AKLEDQ-----LKAVERNVEDYSTEGLEKTTAAKKT 82

DB 360 TLKPRMDRMQGIKAKLEKRLTPDKVKMYDLSLAKKEKEI-----TEKKKLI-AKKS 414

QY 83 ELEKTEADLKKAVNE 97

DB 415 SLKTRGQALKXAVEE 429

RESULT 8

S48396

trpomyosin TPM2 - yeast (*Saccharomyces cerevisiae*)

N;Alternate names: protein YJL138c

C;Species: *Saccharomyces cerevisiae*

C;Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004

C;Accession: S48396; A56490

R;Churcher, C.

submitted to the EMBL Data Library, September 1994

A;Reference number: S48310

A;Accession: S48396

A;Molecule type: DNA

A;Residues: 1-161 <CHU>

A;Cross-references: UNIPROT:P40414; GB:Z47047; EMBL:Z38059; NID:g603997; PID:g763208; MII

J. Cell Biol. 128, 383-392, 1995

A;Title: Tropomyosin is essential in yeast, yet the TPM1 and TPM2 products perform distinct functions

A;Reference number: A56490; MUID:95146545; PMID:7844152

A;Accession: A56490

A;Status: preliminary; nucleic acid sequence not shown

A;Molecule type: DNA

A;Residues: 1-161 <DRE>

A;Cross-references: GB:Z47047; GB:Z38059; NID:g603997; PID:g763208

C;Genetics:

A;Gene: SGD:TPM2

A;Cross-references: SGD:S0001400; MIPS:YJL138c

A;Map position: 9L

C;Superfamily: tropomyosin TPM1

C;Keywords: cytoskeleton

Query Match 22.1%; Score 107.5; DB 2; Length 161;

Best Local Similarity 32.3%; Pred. No. 0.72;

Matches 31; Conservative 21; Mismatches 39; Indels 5; Gaps 2;

QY 6 ESESDYAKGFRAPLHSLDKAKLKLKLELSDKIDELDAEIAKLEDLKXAVEE 61

DB 14 ESESWQEKYEELREQLKEQNTKEIKLSAKNEQDSEVEKLESLDSTKQLAED 73

QY 62 NNVEDYSTEGLEKTTAAKTELEKTEADLKKAVNE 97

DB 74 SNNLRS-NNENTYKKNQDLEQLEDESEAKLKEAMDK 108

RESULT 9

S52536

fcrA 15 protein - Streptococcus pyogenes

C;Species: Streptococcus pyogenes

C;Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004

C;Accession: S52536

R;Katerov, V.; Schalen, C.; Totolian, A.A.

Mol. Gen. Genet. 245, 78-85, 1994

A;Title: Sequencing of genes within the vir regulon of Streptococcus pyogenes type M15 -

A;Reference number: S52535; MUID:95147851; PMID:7845360

A;Accession: S52536

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-388 <KAT>

A;Cross-references: UNIPROT:Q53474; GB:S75411; NID:g914107; PIDN:AAB33261.1; PID:g914109

C;Superfamily: M5 protein

Query Match 21.9%; Score 106.5; DB 2; Length 388;

Best Local Similarity 29.6%; Pred. No. 2;

Matches 37; Conservative 16; Mismatches 39; Indels 33; Gaps 3;

QY 1 LKSIDSESDYAKGFR-----PLHSLDKAKLKLKLE----- 36

DB 171 LNOQDASKTEETAKLQSEAAATLENLGSAKRELTDLQAKLDTATBAKAKLESQVTTLENL 230

QY 37 -----ELSD---KIDELDAEIAKLEDLKXAVEENNVEDYSTEGLEKTTAAKKT 87

DB 231 LGSAKRELTDLQAKLDAANAQKQSQAAALEKQLEATKELADLQAKLAATNQEKEL 290

QY 88 EADLK 92

DB 291 EAEAK 295

RESULT 10

A33939

Fc gamma (IgG) receptor II precursor - Streptococcus sp. (fragment)
C;Species: Streptococcus sp.
C;Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 26-Aug-1999
A;Accession: A33939
R;Heath, D.G.; Cleary, P.P.
Proc. Natl. Acad. Sci. U.S.A. 86, 4741-4745, 1989
A;Title: Fc-receptor and M-protein genes of group A streptococci are products of gene du
A;Reference number: A33939; MUID:89282846; PMID:2660147
A;Accession: A33939
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-405 <HEA>
A;Cross-references: GB:M22532; NID:g153628; PIDN:AA895296.1; PID:g552003
C;Superfamily: M5 protein
C;Keywords: immunoglobulin receptor

Query Match 21.9%; Score 106.5; DB 2; Length 405;
Best Local Similarity 29.6%; Pred. No. 2.1;
Matches 37; Conservative 16; Mismatches 39; Indels 33; Gaps 3;

Qy 1 LKIDESESEDYAKGFRA-----PLHSKLDKQAKLSKLE----- 36
Db 208 LKQDASKTEIAIKLQSEATLENLGSARELTDLQAKLDTATRAEKAKLESQVTTLENL 267
Qy 37 -----ELSD---KIDELDAEIAKLELDQKAVENNVVDYSTEGLEKTTAAKKTLEKT 87
Db 268 LGSARELTDLQAKLDAANAEEKLSQAAALEKLEATKKELADLQAKLAATNQEKEL 327
Qy 88 EADLK 92
Db 328 EAEAK 332

RESULT 11
S70332
outer surface protein F precursor - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S70532
R;Aking, D.R.; Porcella, S.F.; Popova, T.G.; Shevchenko, D.; Baker, S.I.; Li, M.; Norga
Mol. Microbiol. 18, 507-520, 1995
A;Title: Evidence for in vivo but not in vitro expression of a Borrelia burgdorferi oute
A;Reference number: S70531; MUID:96342380; PMID:8748034
A;Accession: S70532
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-229 <AKI>
A;Cross-references: UNIPROT:Q44735; EMBL:U19754; NID:g3318660; PIDN:AAC26147.1; PID:g896
C;Genetics:
A;Gene: ospF
C;Superfamily: outer surface protein F ospF
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-229/Product: outer surface protein F #status predicted <MAT>

Query Match 21.6%; Score 105; DB 2; Length 229;
Best Local Similarity 30.3%; Pred. No. 1.5;
Matches 36; Conservative 27; Mismatches 32; Indels 24; Gaps 7;

Qy 1 LKIDESE-----SEDYAK---EGFRAPLH-----SKLDAKQAKLSKLEELSDKIDELDAE 48
Db 33 VQDLESSEQNVKTEQBIKQKQVEGFLBLETGKLNKLDITKEIE-KRIQELKEKIEKLEAK 91
Qy 49 IAKL-----EDOLKAVEE---NNVVDYSTEGLEKTTAAKKTLEKTEADLKAVNE 97
Db 92 KTSLKTSYSEYEKLUKQIKELKKGADLED-KLKGLEDLSLKXKKEERKKALEDKAKKPEE 149

RESULT 12
T05409
hypothetical protein F10M6.170 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T05409

R;Bevan, M.; Weichselgartner, M.; Fartmann, B.; Granderath, K.; Dauner, D.; Herzi, A.; N;
submitted to the Protein Sequence Database, February 1998
A;Reference number: Z15414
A;Accession: T05409
A;Molecule type: DNA
A;Residues: 1-764 <BEV>
A;Cross-references: UNIPROT:O49371; EMBL:ALQ21811
A;Experimental source: cultivar Columbia; BAC clone F10M6
C;Genetics:
A;Map position: 4
A;Note: F10M6.170

Query Match 21.6%; Score 105; DB 2; Length 764;
Best Local Similarity 30.7%; Pred. No. 5;
Matches 35; Conservative 24; Mismatches 37; Indels 18; Gaps 4;

Qy 2 KEIDSESEDYAKGFRAPLHSLDKDAKQAKLSKL-FELSDKIDELDAEIAKLELDQKAVE 60
Db 163 REIEELKHKLREDEPRAALQSSLTTLKEEELKQROEIANRSEKVSMAISEFESKQLLS 222
Qy 61 ENNV-----EDYS-----TEGLEKTTAAKKTLEK---TEADLKAVNE 97
Db 223 KANENVKQGEIYALQRALEKEEELISKATKLEQEKLRTEANLKKQTEE 276

RESULT 13
S23326
gene M12.2 protein precursor - Streptococcus pyogenes
C;Species: Streptococcus pyogenes
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S23326
R;Bessen, D.E.; Fischetti, V.A.
Infect. Immun. 60, 124-135, 1992
A;Title: Nucleotide sequences of two adjacent M or M-like protein genes of group A strep
A;Reference number: S23325; MUID:92104662; PMID:1370269
A;Accession: S23326
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-372 <BES>
A;Cross-references: UNIPROT:P50469; EMBL:X61276; NID:g47369; PIDN:CAA43582.1; PID:g47371
C;Superfamily: M5 protein

Query Match 21.5%; Score 104.5; DB 2; Length 372;
Best Local Similarity 30.8%; Pred. No. 2.7;
Matches 37; Conservative 17; Mismatches 37; Indels 29; Gaps 5;

Qy 3 EIDSESEDYAK-----EGFRAPLHS-----KLDKQAKLSKLE-----ELSD 40
Db 48 EVKSEKESQYKTLALRGENADLRNVNAKYLEKINAEKKNKLEINKELNENYKLOD 107
Qy 41 KIDELDAEIAKLELDQKAVENNVVDYSTEGLEKTTAAKKT---ELE-----KTEADLKK 93
Db 108 GIDALEKEKEDLTKTLAKTTKENEISEASRKGLSRDLASRTAKKLEAKHQKLEAKNK 167

RESULT 14
S35760
fcrA protein precursor - Streptococcus pyogenes
C;Species: Streptococcus pyogenes
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S35760; A42711
R;Podbielski, A.
submitted to the EMBL Data Library, November 1992
A;Reference number: S35760
A;Accession: S35760
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-415 <POD>
A;Cross-references: UNIPROT:Q54859; EMBL:X69324; NID:g311759; PIDN:CAA49165.1; PID:g3117
R;Haanes, E.J.; Heath, D.G.; Cleary, P.P.
J. Bacteriol. 174, 4967-4976, 1992
A;Title: Architecture of the vir regulons of group A streptococci parallels opacity fac
A;Reference number: A42711; MUID:92332431; PMID:1385809

[illegible]

Search completed: June 21, 2005, 10:11:59

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:54:24 ; Search time 61.3194 Seconds
(without alignments)
826.751 Million cell updates/sec

Title: US-10-674-755-13
Perfect score: 486
Sequence: 1 LKEIDSESEDYAKGFRAP.....KKTELEKTEADLKKA VNEPE 99

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 03: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	474	97.5	415	2 Q9LAY1	Q9lay1 streptococc
2	453	93.2	619	2 Q54972	Q54972 streptococc
3	453	93.2	619	2 Q8DR10	Q8dr10 streptococc
4	446	91.8	417	2 Q9LAY3	Q9lay3 streptococc
5	385	79.2	739	2 Q9ROT4	Q9rot4 streptococc
6	385	79.2	820	2 Q9RQT1	Q9rqt1 streptococc
7	385	79.2	929	2 Q9KK19	Q9kk19 streptococc
8	385	79.2	929	2 Q9ZAY5	Q9zay5 streptococc
9	372	76.5	437	2 Q9LAY4	Q9lay4 streptococc
10	370	76.1	395	2 Q9LAY2	Q9lay2 streptococc
11	370	76.1	408	2 Q9LAY0	Q9lay0 streptococc
12	366	75.3	99	2 Q8KQK4	Q8kk4 streptococc
13	366	75.3	249	2 Q9LS75	Q9ls75 streptococc
14	359	73.9	224	2 Q8GNS8	Q8gns8 streptococc
15	353	72.6	426	2 Q9LAY5	Q9lay5 streptococc
16	350.5	72.1	869	2 Q9KK27	Q9kk27 streptococc
17	339.5	69.9	246	2 Q9LS78	Q9ls78 streptococc
18	338.5	69.7	225	2 Q9LS91	Q9ls91 streptococc
19	334.5	68.8	255	2 Q9LS81	Q9ls81 streptococc
20	334.5	68.8	255	2 Q9LSB6	Q9lsb6 streptococc
21	332.5	68.4	222	2 Q9LS77	Q9ls77 streptococc
22	332.5	68.4	262	2 Q9LS76	Q9ls76 streptococc
23	332.5	68.4	415	2 Q9LAY7	Q9lay7 streptococc
24	329.5	67.8	416	2 Q9LAY8	Q9lay8 streptococc
25	327.5	67.4	406	2 Q9LAZ0	Q9laz0 streptococc
26	324.5	66.8	393	2 Q9LAZ3	Q9laz3 streptococc
27	321.5	66.2	394	2 Q9LAY6	Q9lay6 streptococc
28	321.5	66.2	395	2 Q9LAZ1	Q9laz1 streptococc
29	319.5	65.7	194	2 Q9LSB5	Q9lsb5 streptococc
30	319.5	65.7	218	2 Q6UEB2	Q6ueb2 streptococc
31	319.5	65.7	233	2 Q9LS68	Q9ls68 streptococc

32	319.5	65.7	236	2 Q9LS69	Q9ls69 streptococc
33	319.5	65.7	243	2 Q9LS64	Q9ls64 streptococc
34	319.5	65.7	243	2 Q9LS67	Q9ls67 streptococc
35	319.5	65.7	244	2 Q9LS65	Q9ls65 streptococc
36	319.5	65.7	247	2 Q9LS66	Q9ls66 streptococc
37	319.5	65.7	249	2 Q9LS70	Q9ls70 streptococc
38	319.5	65.7	254	2 Q9LS63	Q9ls63 streptococc
39	319.5	65.7	401	2 Q9LAZ2	Q9laz2 streptococc
40	316.5	65.1	340	2 Q8KQK5	Q8kk5 streptococc
41	313.5	64.5	207	2 Q8GNS9	Q8gns9 streptococc
42	309.5	63.7	237	2 Q9LS92	Q9ls92 streptococc
43	309.5	63.7	395	2 Q9LAY9	Q9lay9 streptococc
44	197.5	40.6	653	2 Q34097	Q34097 streptococc
45	179.5	36.9	246	2 Q9LSB4	Q9lsb4 streptococc

ALIGNMENTS

RESULT 1

Q9LAY1 ID Q9LAY1 PRELIMINARY; PRT; 415 AA.
AC Q9LAY1;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WU2;
RX MEDLINE=20448953; PubMed=10992499;
DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071814; AAF27710.1; -;
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 415 415
SQ SEQUENCE 415 AA; 46075 MW; 213C1AF7FF21642F CRC64;

Query Match 97.5%; Score 474; DB 2; Length 415;
Best Local Similarity 98.0%; Pred. No. 7e-23; Indels 0; Gaps 0;
Matches 97; Conservative 1; Mismatches 1;

QY 1 LKEIDSESEDYAKGFRAPLHSHKLDKQAKLSKLELSDKIDELDAEIAKLEDLQKAVE 60
|||||
Db 221 LKEIDSESEDYAKGFRAPLQSKLDKAKKLSKLELSDKIDELDAEIAKLEDLQKAVE 280
|||||
QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKKA VNEPE 99
|||||
Db 281 ENNVEDYSTEGLEKTIAAKTELEKTEADLKKA VNEPE 319
|||||

RESULT 2

Q54972 ID Q54972 PRELIMINARY; PRT; 619 AA.
AC Q54972;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Pneumococcal surface protein A precursor.
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;

```

RN  SEQUENCE FROM N.A.
RX  MEDLINE=92105030; PubMed=1729249;
RA  Yother J., Briles D.E.;
RT  "Structural properties and evolutionary relationships of PspA, a
RT  surface protein of Streptococcus pneumoniae, as revealed by sequence
RT  analysis.";
RL  J. Bacteriol. 174:601-609(1992).
RN  [2]
RP  SEQUENCE FROM N.A.
RA  Yother J., Briles D.E.;
RL  EMBL; M74122; AAA27018.1; -.
DR  PIR; A41971; A41971.
DR  PIR; A97887; A97887.
DR  HSP; P06653; IHCX.
DR  InterPro; IPR002479; CW binding.
DR  InterPro; IPR002345; Lipocalin.
DR  Pfam; PF01473; CW_binding_1; 10.
DR  PROSITE; PS00213; LIPOCALIN; UNKNOWN_3.
KW  Signal.
FT  SIGNAL 1 31 Potential.
FT  CHAIN 32 619 pneumococcal surface protein A.
SQ  SEQUENCE 619 AA; 68605 MW; 5AA8BDB40C2841CA CRC64;

Query Match 93.2%; Score 453; DB 2; Length 619;
Best Local Similarity 93.9%; Pred. No. 2.2e-21;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAQAKLSKLELSKDIDELDAEIAKLEDLQKAVE 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDAKAKLSKLELSKDIDELDAEIAKLEDLQKAAE 282

Qy 61 ENNVEDYSTEGLEKTIAAKTELEADLKKA VNEPE 99
Db 283 ENNVEDYFKEGLEKTIAAKAELEKTEADLKKA VNEPE 321

RESULT 3
Q8DR10 ID Q8DR10 PRELIMINARY; PRT; 619 AA.
AC Q8DR10;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Surface protein pspA.
GN Names=pspA; OrderedLocustNames=spr0121;
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=171101;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21429245; PubMed=11544234;
RX DOI=10.1128/JB.183.19.5709-5717.2001;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAnnis S.M., McHenney M., McLeaster K., Mundy C.W., Nicas T.I.,
RA Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rockey P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,
RA Glass J.L.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717(2001).
DR EMBL; AE008396; AAK98925.1; -.
DR PIR; A41971; A41971.
DR PIR; A97887; A97887.
DR HSP; P06653; IHCX.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam; PF01473; CW_binding_1; 10.

```

```

DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_3.
SQ SEQUENCE 619 AA; 68605 MW; 5AA8BDB40C2841CA CRC64;

Query Match 93.2%; Score 453; DB 2; Length 619;
Best Local Similarity 93.9%; Pred. No. 2.2e-21;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAQAKLSKLELSKDIDELDAEIAKLEDLQKAVE 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDAKAKLSKLELSKDIDELDAEIAKLEDLQKAAE 282

Qy 61 ENNVEDYSTEGLEKTIAAKTELEADLKKA VNEPE 99
Db 283 ENNVEDYFKEGLEKTIAAKAELEKTEADLKKA VNEPE 321

RESULT 4
Q9LAY3 ID Q9LAY3 PRELIMINARY; PRT; 417 AA.
AC Q9LAY3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE PspA (Fragment).
DE PspA (Fragment).
GN Names=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=EF10197;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071812; AAF27708.1; -.
DR HSP; P00192; 256B.
DR NON_TER 417 417
FT NON_TER 417 417
SQ SEQUENCE 417 AA; 46960 MW; 876EAD3276506BEC CRC64;

Query Match 91.8%; Score 446; DB 2; Length 417;
Best Local Similarity 91.9%; Pred. No. 4.4e-21;
Matches 91; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAQAKLSKLELSKDIDELDAEIAKLEDLQKAVE 60
Db 213 LKEIDSESDYVKEGFRAPLQSELDAAKQAKLSKLELSKDIDELDAEIAKLEDLQKAAE 272

Qy 61 ENNVEDYSTEGLEKTIAAKTELEADLKKA VNEPE 99
Db 273 ENNVEDYFKEGLEKTIAAKAELEKTEADLKKA VNEPE 311

RESULT 5
Q9RQT4 ID Q9RQT4 PRELIMINARY; PRT; 739 AA.
AC Q9RQT4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE Hypothetical protein pspC (Fragment).
DE Hypothetical protein pspC (Fragment).
GN Names=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=E134;

```

```

RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, pspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; AF068647; AAF13457.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW binding_1; 1.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
KW Hypothetical protein.
FT NON TER 739
SQ SEQUENCE 739 AA; 83960 MW; 7EE2F2F676ABF989 CRC64;

Query Match 79.2%; Score 385; DB 2; Length 739;
Best Local Similarity 81.8%; Pred. No. 5.9e-17;
Matches 81; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 LXEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQKAVE 60
DB 537 LXEIDSESDYKLEGLRAPLQSKLDTKKAQKLSKLELSDKIDELDAETAKLEVQLKDAE 596
QY 61 ENNVNVEDYSTEGLEKTTAAKTELEKTEADLKAVNEPE 99
DB 597 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKAVDEPE 635

RESULT 6
Q9RQT1 ID Q9RQT1 PRELIMINARY; PRT; 820 AA.
AC Q9RQT1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein pspC (Fragment).
GN Names=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9183;
RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, pspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; AF068650; AAF13460.1; -.
DR HSSP; P04268; IIC2.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW binding_1; 1.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
KW Hypothetical protein.
FT NON TER 820
SQ SEQUENCE 820 AA; 91752 MW; 33C095849ABB0942 CRC64;

Query Match 79.2%; Score 385; DB 2; Length 820;
Best Local Similarity 81.8%; Pred. No. 6.5e-17;
Matches 81; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

```

```

QY 1 LXEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQKAVE 60
DB 530 LXEIDSESDYKLEGLRAPLQSKLDTKKAQKLSKLELSDKIDELDAETAKLEVQLKDAE 589
QY 61 ENNVNVEDYSTEGLEKTTAAKTELEKTEADLKAVNEPE 99
DB 590 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKAVDEPE 628

RESULT 7
Q9KK19 ID Q9KK19 PRELIMINARY; PRT; 929 AA.
AC Q9KK19;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein PspC.
GN Names=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=srf10;
RX MEDLINE=21888621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae.";
RL Gene 284:63-71(2002).
DR EMBL; AF154037; AAF73809.1; -.
DR HSSP; P06653; IH8G.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
SQ SEQUENCE 929 AA; 105003 MW; 2DC8293302FAFA64 CRC64;

Query Match 79.2%; Score 385; DB 2; Length 929;
Best Local Similarity 81.8%; Pred. No. 7.3e-17;
Matches 81; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 LXEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQKAVE 60
DB 530 LXEIDSESDYKLEGLRAPLQSKLDTKKAQKLSKLELSDKIDELDAETAKLEVQLKDAE 589
QY 61 ENNVNVEDYSTEGLEKTTAAKTELEKTEADLKAVNEPE 99
DB 590 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKAVDEPE 628

RESULT 8
Q9ZAY5 ID Q9ZAY5 PRELIMINARY; PRT; 929 AA.
AC Q9ZAY5;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein C.
GN Names=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF6796;

```

```

RX MEDLINE=20038319; PubMed=10569772;
RA Brooke-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
protein, PspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; U72655; AAD00184.1; -.
DR HS3P; P06653; 1HCX.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
SQ SEQUENCE 929 AA; 104991 MW; 2DC8293302FFB081 CRC64;

Query Match 79.2%; Score 385; DB 2; Length 929;
Best Local Similarity 81.8%; Pred. No. 7.3e-17;
Matches 81; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKIDELDAEIAKLELDQKAVE 60
Db 530 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKIDELDAEIAKLELDQKAVE 589
Qy 61 ENNVEDYSTEGLEKTTAAKTELEKTEADLKKAVNEPE 99
Db 590 GNNVVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 628

RESULT 9
Q9LAY4
ID Q9LAY4 PRELIMINARY; PRT; 437 AA.
AC Q9LAY4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E134;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071813; AAF27709.1; -.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin alc DH like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYSIN.
FT NON TER 437
SQ SEQUENCE 437 AA; 48071 MW; F5EFD2CD13E08CD8 CRC64;

Query Match 76.5%; Score 372; DB 2; Length 437;
Best Local Similarity 79.8%; Pred. No. 2.5e-16;
Matches 79; Conservative 3; Mismatches 17; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKIDELDAEIAKLELDQKAVE 60
Db 235 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKIDELDAEIAKHVVQLKDAE 294
Qy 61 ENNVEDYSTEGLEKTTAAKTELEKTEADLKKAVNEPE 99
Db 295 GNNVVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 333

us-10-674-755-13.rup

```

```

RESULT 10
Q9LAY2
ID Q9LAY2 PRELIMINARY; PRT; 395 AA.
AC Q9LAY2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E56796;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071813; AAF27709.1; -.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin alc DH like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYSIN.
FT NON TER 395
SQ SEQUENCE 395 AA; 42963 MW; 58E6EF956BCBCC1E CRC64;

Query Match 76.1%; Score 370; DB 2; Length 395;
Best Local Similarity 78.8%; Pred. No. 3.1e-16;
Matches 78; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKIDELDAEIAKLELDQKAVE 60
Db 225 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKIDELDAEIAELEVQLKDAE 284
Qy 61 ENNVEDYSTEGLEKTTAAKTELEKTEADLKKAVNEPE 99
Db 285 GNNVVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 323

RESULT 11
Q9LAY0
ID Q9LAY0 PRELIMINARY; PRT; 408 AA.
AC Q9LAY0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9163;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071815; AAF27711.1; -.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin alc DH like.
FT NON TER 408
SQ SEQUENCE 408 AA; 44254 MW; 4F64D874217297EF CRC64;

Query Match 76.1%; Score 370; DB 2; Length 408;
Best Local Similarity 78.8%; Pred. No. 3.2e-16;

```

```

RESULT 13
Q9L575 PRELIMINARY; PRT; 249 AA.
ID AC Q9L575;
AC AC Q9L575;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PSpA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=195;
EX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
FT pneumococcal strains in the United States and of internationally

```

RESULT 15	
Q9LAY5	
ID Q9LAYS	PRELIMINARY; PRT; 426 AA.
AC Q9LAYS;	
DT 01-OCT-2000	(TREMBlrel. 15, Created)
DT 01-OCT-2000	(TREMBlrel. 15, Last sequence update)

```
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBL5;
RX MEDLINE=20448953; PubMed=10992499;
RA DOI=10.1128/IAI.68.10.5889-5900.2000; D.E.;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RL in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071810; AAP27706.1; -.
DR HSP; P00192; IMGT.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 426
SQ SEQUENCE 426 AA; 46534 MW; 81AA11348CBE6634 CRC64;

Query Match 72.6%; Score 353; DB 2; Length 426;
Best Local Similarity 74.7%; Pred. No. 4e-15;
Matches 74; Conservative 8; Mismatches 17; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAQAKLSKLELSKIDELDAEIAKLELDQKAVE 60
Db 215 LKQINESDSEDYVKEGLRAPLQSELDTKAKLLKLELSKIEELDAEIAELEVQLKDAE 274

Qy 61 ENNVVEDYSTGLEKTIAAKTELEKTEADLKKA VNEPE 99
Db 275 GNNVVEAYFKEGLEKTTAEKKAELKABADLKKA VDEPE 313
```

Search completed: June 21, 2005, 10:22:11
Job time : 62.3194 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:50:39 ; Search time 73.8459 Seconds
(without alignments)
518.502 Million cell updates/sec

Title: US-10-674-755-14

Perfect score: 482

Sequence: 1 LKDIIDSESDYAKGERAP.....KKAELEKARADLKKAVIDPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	472	97.9	206	2 AAW14574	Aaw14574 Streptoco
2	472	97.9	206	7 ABW02608	Abw02608 Db15c pne
3	472	97.9	8991	6 ABU08487	Abu08487 S. pneumo
4	440	91.3	170	7 ABW02614	Abw02614 Rct135c p
5	440	91.3	181	7 ABW02596	Abw02596 0922134c
6	440	91.3	865	6 ABU08489	Abu08489 S. pneumo
7	440	91.3	929	2 AAW14593	Aaw14593 Streptoco
8	440	91.3	929	2 AAY43384	Aay43384 S. pneumo
9	437	90.7	188	2 AAW14580	Aaw14580 Streptoco
10	437	90.7	188	7 ABW02613	Abw02613 Rct129c p
11	428	88.8	588	6 ABU08491	Abu08491 Coiled co
12	428	88.8	589	2 AAY43392	Aay43392 PspC alph
13	426	88.4	204	2 AAW14578	Aaw14578 Streptoco
14	426	88.4	204	7 ABW02612	Abw02612 Rct123c p
15	425.5	88.3	180	2 AAW14562	Aaw14562 Streptoco
16	422.5	87.7	187	2 AAW14579	Aaw14579 Streptoco
17	405	84.0	1231	6 ABU08490	Abu08490 Fragment
18	389	80.7	198	2 AAW14581	Aaw14581 Streptoco
19	386	80.1	198	7 ABW02615	Abw02615 Rxic pneu
20	386	80.1	204	2 AAW14571	Aaw14571 Streptoco
21	386	80.1	204	7 ABW02605	Abw02605 Bf1019c p
22	386	80.1	315	2 AAY04375	Aay04375 Streptoco
23	386	80.1	619	2 AAR63437	Aar63437 Pneumococ
24	386	80.1	619	2 AAR87598	Aar87598 Pneumococ
25	386	80.1	619	2 AAR86911	Aar86911 Pneumococ

ALIGNMENTS

RESULT 1

AAW14574

ID AAW14574 standard; protein; 206 AA.

XX AAW14574;

AC AAW14574;

DT 17-OCT-2003 (revised)

DT 28-OCT-1997 (first entry)

XX Streptococcus pneumoniae PspA central region.

DE PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;

KW bacteraemia; pneumonia.

XX Streptococcus pneumoniae; strain Db15.

OS Streptococcus pneumoniae; strain Db15.

XX Key Location/Qualifiers

FH Misc-difference 50

FT /note= "unidentified amino acid"

XX WO9709994-A1.

PN 20-MAR-1997.

XX 16-SEP-1996; 96WO-US014819.

XX 15-SEP-1995; 95US-00529055.

XX (UABR-) UAB RES FOUND.

PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;

PI Hollingshead S, Tart R, Brooks-Walter A;

XX WPI; 1997-202002/18.

XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used

XX in vaccines for protecting animals against S.pneumoniae infection.

PS Example 6; Fig 13; 296pp; English.

XX This sequence shows the central portion, including the C-terminus of the

CC alpha-helix region and some of the proline-rich region, of pneumococcal

CC surface protein A (PspA) of Streptococcus pneumoniae strain Db15.

CC Comparison of the N-terminal and central regions (AAW14533-57 and

CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can

CC be used to divide the strains into several families based on sequence

CC homologies. PspA polypeptides, or fragments of them, can be used in

CC vaccines to protect animals against S. pneumoniae infection and hence for

CC the prevention of diseases such as otitis media, meningitis, bacteraemia
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
 CC region and the immediate 5' tip of the coding sequence are likely to be
 CC the critical sequences for predicting PspA cross-reactions and vaccine
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)

XX
 SQ Sequence 206 AA;

Query Match 97.9%; Score 472; DB 2; Length 206;
 Best Local Similarity 99.0%; Pred. No. 2.2e-37;
 Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKIDSDSDYAKGGRAPLQSELDTKKAKLLKLELSGKIELDAEIXEVLQKDAE 60
 |||||
 Db 1 LKIDSDSDYAKGGRAPLQSELDTKKAKLLKLELSGKIELDAEIXEVLQKDAE 60
 |||||
 Qy 61 GNNVAYFKGEGLEKTTAEKKAELKAEADLKKAVDEPE 99
 |||||
 Db 61 GNNVAYFKGEGLEKTTAEKKAELKAEADLKKAVDEPE 99
 |||||

RESULT 2
 ABW02608
 ID ABW02608 standard; protein; 206 AA.
 XX
 AC ABW02608;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 DE Db15c pneumococcal surface protein A (PspA) central region.
 XX
 KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
 KW immunological; gene therapy; immunostimulant.
 XX
 OS Unidentified.

XX
 FH Key Location/Qualifiers
 FT Misc-difference 1. .206
 FT /note= "Xaa = Unknown amino acid"
 XX
 XX US6592876-B1.
 XX
 PD 15-JUL-2003.
 XX
 XX 15-SEP-1995; 95US-00529055.
 XX
 XX 20-APR-1993; 93US-00048896.
 PR
 PR 06-JUN-1995; 95US-00465746.
 XX
 XX (UABR-) UAB RES FOUND.
 XX
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
 PI
 PI WPI; 2003-862841/80.
 DR
 DR
 XX
 XX Immunological composition for obtaining expression products used for
 PT detecting the presence of Streptococcus pneumoniae or its strain,
 PT comprises at least two different full length isolated gene encoding
 PT pneumococcal surface protein A.
 XX
 XX
 PS Example 6; SEQ ID NO 54; 121pp; English.

XX
 CC The present invention relates to an immunological composition comprising
 CC at least 2 different full length isolated genes encoding pneumococcal
 CC surface protein A (PspAs) from different groups based on restriction
 CC fragment polymorphism analysis. The invention is useful for obtaining
 CC expression products by recombinant techniques to detect, determine,
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its
 CC strain. The expression product is useful for preparing antigenic,
 CC immunological or vaccine compositions, for eliciting antibodies, an
 CC immunological response (other than or additional to antibodies) or a
 CC protective response (including antibody or other immunological response
 CC by administering compositions to a host). The invention is also useful as

CC vaccines and in gene therapy. The present sequence is Db15c pneumococcal
 CC surface protein A (PspA) central region. This sequence is used in the
 CC exemplification of the invention

XX
 SQ Sequence 206 AA;

Query Match 97.9%; Score 472; DB 7; Length 206;
 Best Local Similarity 99.0%; Pred. No. 2.2e-37;
 Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKIDSDSDYAKGGRAPLQSELDTKKAKLLKLELSGKIELDAEIXEVLQKDAE 60
 |||||
 Db 1 LKIDSDSDYAKGGRAPLQSELDTKKAKLLKLELSGKIELDAEIXEVLQKDAE 60
 |||||
 Qy 61 GNNVAYFKGEGLEKTTAEKKAELKAEADLKKAVDEPE 99
 |||||
 Db 61 GNNVAYFKGEGLEKTTAEKKAELKAEADLKKAVDEPE 99
 |||||

RESULT 3
 ABU08487
 ID ABU08487 standard; protein; 8991 AA.
 XX
 AC ABU08487;
 XX
 DT 24-JUN-2003 (first entry)
 XX
 DE S. pneumoniae pneumococcal surface protein A (PspA) protein.
 XX
 KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
 KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
 KW antibacterial.
 XX
 OS Streptococcus pneumoniae.

XX
 FH Key Location/Qualifiers
 FT Misc-difference 1. .8991
 FT /note= "All Xaa residues within this sequence are
 FT unknown"
 XX
 XX US6500613-B1.
 XX
 PD 31-DEC-2002.
 XX
 XX 16-SEP-1996; 96US-00714741.
 XX
 XX 15-SEP-1995; 95US-00529055.
 PR
 XX (UYAL-) UNIV ALABAMA.
 XX
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
 PI
 PI Hollingshead S, Tart R, Brooks-Walter A;
 XX
 XX WPI; 2003-361534/34.
 DR
 DR
 XX
 XX Isolated PspC amino acid sequence used as polymerase chain reaction or
 PT hybridization probe, comprises pneumococcal surface protein having alpha-
 PT helical, proline rich and repeat regions.
 XX
 XX Disclosure; Col 145-188; 186pp; English.

XX
 CC The present invention relates to the isolation of Streptococcus
 CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
 CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA) -
 CC like protein having alpha-helical, proline rich and repeat regions. The
 CC PspC and PspA proteins may be used in a vaccine to protect against
 CC pneumococcal infections. The polynucleotide sequences encoding PspC and
 CC PspA may be used for the expression of the proteins, and as PCR primers
 CC or hybridisation probes. The present sequence represents S. pneumoniae
 CC PspA protein

XX
 SQ Sequence 8991 AA;

Db	1	LKEIDSDSDYLKEGLRAPLOSQKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLQKDAE	6
Qy	61	GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE	99
Db	61	GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE	99
RESULT 5			
ABW02596			
ID	ABW02596	standard; protein; 181 AA.	
XX	AC	ABW02596;	
XX	AC		
XX	DT	12-FEB-2004 (first entry)	
XX	XX		
DE	0922134c	pneumococcal surface protein A (PspA) central region.	
XX			
KW	Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;		
KW	immunological; gene therapy; immunostimulant.		
XX	Unidentified.		
OS	OS		
XX	US6592876-B1.		
PN	XX		
XX	15-JUL-2003.		
PD	XX		
XX	15-SEP-1995;	95US-00529055.	
XX	XX		
PR	20-APR-1993;	93US-00048896.	
PR	06-JUN-1995;	95US-00465746.	
XX	(UABR-) UAB RES FOUND.		
XX	XX		
PI	Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;		
XX	WPI; 2003-862841/80.		
DR	XX		
XX			
PT	Immunological composition for obtaining expression products used for		
PT	detecting the presence of Streptococcus pneumoniae or its strain,		
PT	comprises at least two different full length isolated gene encoding		
PT	pneumococcal surface protein A.		
XX			
XX	Example 6; SEQ ID NO 42; 121pp; English.		
PS	XX		
CC	The present invention relates to an immunological composition comprising		
CC	at least 2 different full length isolated genes encoding pneumococcal		
CC	surface protein A (PspAs) from different groups based on restriction		
CC	fragment polymorphism analysis. The invention is useful for obtaining		
CC	expression products by recombinant techniques to detect, determine,		
CC	isolate or diagnose the presence of Streptococcus pneumoniae or its		
CC	strain. The expression product is useful for preparing antigenic,		
CC	immunological or vaccine compositions, for eliciting antibodies, an		
CC	immunological response (other than or additional to antibodies) or a		
CC	protective response (including antibody or other immunological response		
CC	by administering compositions to a host). The invention is also useful a		
CC	vaccines and in gene therapy. The present sequence is 0922134c		
CC	pneumococcal surface protein A (PspA) central region. This sequence is		
CC	used in the exemplification of the invention		
XX			
XX	Seq	Sequence 181 AA;	
Query Match 91.3%; Score 440; DB 7; Length 181;			
Best Local Similarity 90.9%; Pred. No. 2.2e-34;			
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps			
Qy	1	LKDIDSDSDYAKGERAPLQSELDTKKAKLLKLEELSGKIEELDAEIAKLEVLQKDAE	60
Db	1	LKEIDSDSDYLKEGLRAPLOSQKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLQKDAE	60
Qy	61	GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE	99
Db	61	GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE	99

```

RESULT 6
ABU08489
ID ABU08489 standard; protein; 865 AA.
XX AC ABU08489;
XX DT 24-JUN-2003 (first entry)
XX DE S. pneumoniae pneumococcal surface protein C (PspC) protein.
XX KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
XX KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
XX KW antibacterial.
XX OS Streptococcus pneumoniae.
XX FH Key Location/Qualifiers
XX FT Peptide 1..37
XX FT Protein /label= Signal_peptide
XX FT /label= Mature_PspC_protein
XX US6500613-B1.
XX 31-DEC-2002. 96US-00714741.
XX 16-SEP-1996; 96US-00714741.
XX 15-SEP-1995; 95US-00529055.
XX (UVAL-) UNIV ALABAMA.
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
XX Hollingshead S, Tart R, Brooks-Walter A;
XX WPI; 2003-361534/34.
XX N-PSDB; ABX95377.
XX Isolated PspC amino acid sequence used as polymerase chain reaction or
XX hybridization probe, comprises pneumococcal surface protein having alpha-
XX helical, proline rich and repeat regions.
XX Claim 3; Fig 21; 186pp; English.
XX The present invention relates to the isolation of Streptococcus
XX pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
XX sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-
XX like protein having alpha-helical, proline rich and repeat regions. The
XX PspC and PspA proteins may be used in a vaccine to protect against
XX pneumococcal infections. The polynucleotide sequences encoding PspC and
XX PspA may be used for the expression of the proteins, and as PCR primers
XX or hybridisation probes. The present sequence represents S. pneumoniae
XX PspC protein
XX SQ Sequence 865 AA;
Query Match 91.3%; Score 440; DB 6; Length 865;
Best Local Similarity 90.9%; Pred. No. 1.5e-33;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
Qy 1 LKQIDSDSDYAKGGERAPLQSLDTKKAKLLKLEELSGKIELDRAEIXEVLQKDAE 60
Db 466 LKEIDSDSDYKLEGLRAPLQSLDTKKAKLLKLEELSDKIDELDRAEIXEVLQKDAE 525
Qy 61 GNNVVEYFKGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 526 GNNVVEYFKGLEKTTAEKKAELKAEADLKKAVDEPE 564
RESULT 7
AAW14593
ID AAW14593 standard; protein; 929 AA.

```

```

XX AC AAW14593;
XX DT 17-OCT-2003 (revised)
XX DT 27-OCT-1997 (first entry)
XX DE Streptococcus pneumoniae PspC.
XX KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
XX KW bacteraemia; pneumonia.
XX OS Streptococcus pneumoniae; strain EF6796.
XX FH Key Location/Qualifiers
XX FT Peptide 1..37
XX FT Protein /label= Sig_peptide
XX FT /label= Mat_protein
XX FT Domain 38..529
XX FT /label= Alpha-helix
XX FT Region 41..242
XX FT /label= Repeat_1
XX FT /note= "alpha-helical repeat region"
XX FT Region 69..637
XX FT /label= Coiled-coil
XX FT /note= "breaks in 7-residue periodicity of the coiled-
XX FT coil occur at amino acids 136-141, 261-304 and 383-387"
XX FT Region 332..492
XX FT /label= Repeat_2
XX FT /note= "alpha-helical repeat region"
XX FT Domain 627..689
XX FT /label= Proline-rich
XX FT Domain 913..929
XX FT /label= C-terminal
XX WO9709994-A1.
XX 20-MAR-1997.
XX 16-SEP-1996; 96WO-US014819.
XX 15-SEP-1995; 95US-00529055.
XX (UABR-) UAB RES FOUND.
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
XX Hollingshead S, Tart R, Brooks-Walter A;
XX WPI; 1997-202002/18.
XX N-PSDB; AAT61728.
XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
XX in vaccines for protecting animals against S.pneumoniae infection.
XX PS Disclosure; Fig 13; 296pp; English.
XX This sequence comprises the pneumococcal protein surface C (pspc) of
XX Streptococcus pneumoniae strain EF6796. The sequence was deduced from the
XX pspC gene (AAT61728). Like PspA, PspC has an alpha-helical coiled-coil
XX region, proline-rich central region, repeat regions, with a choline
XX binding motif, and a C-terminal 17-amino acid tail. The 2 polypeptides
XX share 3 regions of high sequence identity. One is a protection-eliciting
XX region present within the alpha-helical domain. The others are the
XX proline-rich domain and a repeat domain shared with other choline-binding
XX proteins and thought to play a role in cell surface association. PspC and
XX PspA polypeptides, and their fragments, can be used in vaccines to
XX protect against S. pneumoniae infection and hence for the prevention of
XX diseases such as otitis media, meningitis, bacteraemia and pneumonia.
XX (Updated on 17-OCT-2003 to standardise OS field)
XX SQ Sequence 929 AA;
Query Match 91.3%; Score 440; DB 2; Length 929;

```

Best Local Similarity 90.9%; Pred. No. 1.6e-33; Mismatches 4; Indels 0; Gaps 0;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKIDSDSDYAKGERAPLOSELDTKKAKLKLKLELSGKIEELDAEIXELEVQLKDAE 60
DB 530 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLELSGKIEELDAEIXELEVQLKDAE 589

QY 61 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99
DB 590 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 628

RESULT 8
AY43384
ID AAY43384 standard; protein; 929 AA.
XX
AC AAY43384;
XX
DT 27-JAN-2000 (first entry)
XX
DE S. pneumoniae PspC protein sequence.
XX
KW PspC gene; pneumococcal surface protein C; epitope; diagnosis;
KW epitopic region; immunogenic composition; vaccine composition; therapy;
KW meningitis; immunological response; otitis media; bacteraemia; pneumonia;
KW anti-PspA antibody.
XX
OS Streptococcus pneumoniae.
XX
PN WO9953940-A1.
XX
PD 28-OCT-1999.
XX
PF 23-APR-1999; 99WO-US008895.
XX
PR 23-APR-1998; 98US-0082728P.
XX
PA (UYAL-) UNIV ALABAMA.
XX
PI Briles DE, Hollingshead SK, Brooks-Walter A;
DR WPI; 1999-620581/53.
DR N-PSDB; AAZ31956.
XX
XX New epitope polypeptides of Pneumococcal surface protein C, used to
PT develop products for immunological, immunogenic or vaccine compositions,
PT particularly against Streptococcus pneumoniae infections.
XX
PS Example; Fig 11; 109pp; English.
XX
CC This sequence is the Streptococcus pneumoniae pneumococcal surface
CC protein C. The invention relates to an isolated and/or purified
CC polypeptide comprising at least one epitope or epitopic region of
CC pneumococcal surface protein C (PspC). The polypeptides or vectors
CC containing sequence encoding them can be used as immunogenic,
CC immunological or vaccine compositions. The compositions can be used for
CC eliciting an immunological response against Streptococcus pneumoniae
CC (SP), which can cause otitis media, meningitis, bacteraemia and
CC pneumonia. They can be used for eliciting an anti-PspA antibody. The
CC nucleic acid molecules can also be used for detecting pspC, pspA or SP
XX
SQ Sequence 929 AA;

Query Match 91.3%; Score 440; DB 2; Length 929;
Best Local Similarity 90.9%; Pred. No. 1.6e-33;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKIDSDSDYAKGERAPLOSELDTKKAKLKLKLELSGKIEELDAEIXELEVQLKDAE 60
DB 530 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLELSGKIEELDAEIXELEVQLKDAE 589

QY 61 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99
DB 590 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 628

Best Local Similarity 90.9%; Pred. No. 1.6e-33; Mismatches 4; Indels 0; Gaps 0;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKIDSDSDYAKGERAPLOSELDTKKAKLKLKLELSGKIEELDAEIXELEVQLKDAE 60
DB 530 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLELSGKIEELDAEIXELEVQLKDAE 589

QY 61 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99
DB 590 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 628

Db 590 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 628

RESULT 9
AAW14580
ID AAW14580 standard; protein; 188 AA.
XX
AC AAW14580;
XX
DT 17-OCT-2003 (revised)
DT 28-OCT-1997 (first entry)
XX
DE Streptococcus pneumoniae PspA central region.
XX
KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW bacteraemia; pneumonia.
XX
OS Streptococcus pneumoniae; strain Rct135.
XX
PN WO9709994-A1.
XX
PD 20-MAR-1997.
XX
PF 16-SEP-1996; 96WO-US014819.
XX
PR 15-SEP-1995; 95US-00529055.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX
DR WPI; 1997-202002/18.
XX
PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used
PT in vaccines for protecting animals against S.pneumoniae infection.
XX
PS Example 6; Fig 13; 296pp; English.
XX
CC This sequence shows the central portion, including the C-terminus of the
CC alpha-helix region and some of the proline-rich region, of pneumococcal
CC surface protein A (PspA) of Streptococcus pneumoniae strain Rct135.
CC Comparison of the N-terminal and central regions (AAW14533-57 and
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
CC be used to divide the strains into several families based on sequence
CC homologies. PspA polypeptides, or fragments of them, can be used in
CC vaccines to protect animals against S. pneumoniae infection and hence for
CC the prevention of diseases such as otitis media, meningitis, bacteraemia
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
CC region and the immediate 5' tip of the coding sequence are likely to be
CC the critical sequences for predicting PspA cross-reactions and vaccine
CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 188 AA;

Query Match 90.7%; Score 437; DB 2; Length 188;
Best Local Similarity 89.9%; Pred. No. 4.5e-34;
Matches 89; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKIDSDSDYAKGERAPLOSELDTKKAKLKLKLELSGKIEELDAEIXELEVQLKDAE 60
DB 530 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLELSGKIEELDAEIXELEVQLKDAE 60

QY 61 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99
DB 590 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 628

RESULT 10
ABW02613
ID ABW02613 standard; protein; 188 AA.
XX
AC ABW02613;

XX 12-FEB-2004 (first entry)
 XX Rct129c pneumococcal surface protein A (PspA) central region.
 DE
 XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
 KW immunological; gene therapy; immunostimulant.
 KW
 XX Unidentified.
 OS
 XX US6592876-B1.
 PN
 XX 15-JUL-2003.
 PD
 XX 15-SEP-1995; 95US-00529055.
 XX
 PF 20-APR-1993; 93US-00048896.
 XX
 PR 06-JUN-1995; 95US-00465746.
 XX
 XX (UABR-) UAB RES FOUND.
 PA
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
 PI
 XX WPI; 2003-862841/80.
 DR
 XX Immunological composition for obtaining expression products used for
 PT detecting the presence of Streptococcus pneumoniae or its strain.
 PT comprises at least two different full length isolated gene encoding
 PT pneumococcal surface protein A.
 XX
 XX Example 6; SEQ ID NO 59; 121pp; English.
 PS
 XX The present invention relates to an immunological composition comprising
 CC at least 2 different full length isolated genes encoding pneumococcal
 CC surface protein A (pspas) from different groups based on restriction
 CC fragment polymorphism analysis. The invention is useful for obtaining
 CC expression products by recombinant techniques to detect, determine,
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its
 CC strain. The expression product is useful for preparing antigenic,
 CC immunological or vaccine compositions, for eliciting antibodies, an
 CC immunological response (other than or additional to antibodies) or a
 CC protective response (including antibody or other immunological response
 CC by administering compositions to a host). The invention is also useful as
 CC vaccines and in gene therapy. The present sequence is Rct129c
 CC pneumococcal surface protein A (PspA) central region. This sequence is
 CC used in the exemplification of the invention
 XX
 SQ Sequence 188 AA;
 Query Match 90.7%; Score 437; DB 7; Length 188;
 Best Local Similarity 89.9%; Pred. No. 4.5e-34;
 Matches 89; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
 QY 1 LKIDSDSDYAKGERAPLQSELDTPKAKLKLSELSGKIBELDAEIXEVLQKDAE 60
 DB 1 LKIDSDSDYAKGERAPLQSELDTPKAKLKLSELSGKIBELDAEIXEVLQKDAE 60
 QY 61 GNNVVEYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
 DB 61 GNNVVEYFKEGLEKTTAEKKAELKAEADLKKAVDEPD 99
 RESULT 11
 ABU08491
 ID ABU08491 standard; protein; 588 AA.
 XX
 AC ABU08491;
 XX
 XX 24-JUN-2003 (first entry)
 DT
 DE Coiled coil motif of alpha-helix of S. pneumoniae PspC protein.
 XX
 KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;

KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
 XX antibacterial.
 XX Streptococcus pneumoniae.
 OS
 XX US6500613-B1.
 PN
 XX 31-DEC-2002.
 PD
 XX 16-SEP-1996; 96US-00714741.
 XX
 PF 15-SEP-1995; 95US-00529055.
 XX
 PR (UYAL-) UNIV ALABAMA.
 XX
 PA Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
 PI Hollingshead S, Tart R, Brooks-Walter A;
 XX
 XX WPI; 2003-361534/34.
 DR
 XX Isolated PspC amino acid sequence used as polymerase chain reaction or
 PT hybridization probe, comprises pneumococcal surface protein having alpha-
 PT helical, proline rich and repeat regions.
 PT
 XX Disclosure; Fig 23; 186pp; English.
 PS
 XX The present invention relates to the isolation of Streptococcus
 CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
 CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-
 CC like protein having alpha-helical, proline rich and repeat regions. The
 CC PspC and PspA proteins may be used in a vaccine to protect against
 CC pneumococcal infections. The polynucleotide sequences encoding PspC and
 CC PspA may be used for the expression of the proteins, and as PCR primers
 CC or hybridisation probes. The present sequence represents a coiled coil
 CC motif of the alpha-helix of S. pneumoniae PspC protein
 XX
 SQ Sequence 588 AA;
 Query Match 88.8%; Score 428; DB 6; Length 588;
 Best Local Similarity 90.7%; Pred. No. 1.3e-32;
 Matches 88; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 QY 1 LKIDSDSDYAKGERAPLQSELDTPKAKLKLSELSGKIBELDAEIXEVLQKDAE 60
 DB 492 LKIDSDSDYAKGERAPLQSELDTPKAKLKLSELSGKIBELDAEIXEVLQKDAE 551
 QY 61 GNNVVEYFKEGLEKTTAEKKAELKAEADLKKAVDE 97
 DB 552 GNNVVEYFKEGLEKTTAEKKAELKAEADLKKAVDE 588
 RESULT 12
 AAY43392
 ID AAY43392 standard; protein; 589 AA.
 XX
 AC AAY43392;
 XX
 XX 27-JAN-2000 (first entry)
 DT
 XX PspC alpha-helix coiled-coil region.
 DE
 XX PspC gene; pneumococcal surface protein C; epitope; diagnosis;
 KW epitopic region; immunogenic composition; vaccine composition; therapy;
 KW meningitis; immunological response; otitis media; bacteraemia; pneumonia;
 KW anti-PspA antibody.
 XX
 XX Streptococcus pneumoniae.
 OS
 XX WO9953940-A1.
 PN
 XX 28-OCT-1999.
 PD
 XX 23-APR-1999; 99WO-US008895.
 PF

XX 23-APR-1998; 98US-0082728P.
 XX (UYAL-) UNIV ALABAMA.
 XX Briles DE, Hollingshead SK, Brooks-Walter A;
 XX WPI; 1999-620581/53.
 XX New epitope polypeptides of Pneumococcal surface protein C, used to
 PT develop products for immunological, immunogenic or vaccine compositions,
 PT particularly against Streptococcus pneumoniae infections.
 XX Example 1; Fig 3; 109pp; English.
 XX This sequence is the coiled-coil region of the Streptococcus pneumoniae
 CC pneumococcal surface protein C. The invention relates to an isolated
 CC and/or purified polypeptide comprising at least one epitope or epitopic
 CC region of Pneumococcal surface protein C (PspC). The polypeptides or
 CC vectors containing sequence encoding them can be used as immunogenic,
 CC immunological or vaccine compositions. The compositions can be used for
 CC eliciting an immunological response against Streptococcus pneumoniae
 CC (SP), which can cause otitis media, meningitis, bacteraemia and
 CC pneumonia. They can be used for eliciting an anti-PspA antibody. The
 CC nucleic acid molecules can also be used for detecting pspC, pspA or SP
 XX
 SQ Sequence 589 AA;
 Query Match 88.8%; Score 428; DB 2; Length 589;
 Best Local Similarity 90.7%; Pred. No. 1.3e-32;
 Matches 88; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 QY 1 LKIDSDSDYAKGERAPLOSELDTKKAKLKLKLELSGKIELDAEIXEVLKDAE 60
 Db 493 LKIDSDSDYAKGERAPLOSELDTKKAKLKLKLELSGKIELDAEIXEVLKDAE 60
 QY 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDE 97
 Db 553 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDE 589
 RESULT 13
 AAW14578
 ID AAW14578 standard; protein; 204 AA.
 XX
 AC AAW14578;
 XX
 DT 17-OCT-2003 (revised)
 DE 28-OCT-1997 (first entry)
 XX
 DE Streptococcus pneumoniae PspA central region.
 XX
 KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
 KW bacteraemia; pneumonia.
 XX
 OS Streptococcus pneumoniae; strain Rct123.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 4 /note= "unidentified amino acid"
 FT Misc-difference 8 /note= "unidentified amino acid"
 FT Misc-difference 8 /note= "unidentified amino acid"
 XX
 XX WO9705994-A1.
 XX
 XX 20-MAR-1997.
 XX
 XX 16-SEP-1996; 96WO-US014819.
 XX
 XX 15-SEP-1995; 95US-00529055.
 XX
 XX (UABR-) UAB RES FOUND.
 XX

PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
 XX Hollingshead S, Tart R, Brooks-Walter A;
 XX WPI; 1997-202002/18.
 XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
 PT in vaccines for protecting animals against S.pneumoniae infection.
 XX
 PS Example 6; Fig 13; 296pp; English.
 XX This sequence shows the central portion, including the C-terminus of the
 CC alpha-helix region and some of the proline-rich region, of pneumococcal
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Rct123.
 CC Comparison of the N-terminal and central regions (AAW14533-57 and
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
 CC be used to divide the strains into several families based on sequence
 CC homologies. PspA polypeptides, or fragments of them, can be used in
 CC vaccines to protect animals against S. pneumoniae infection and hence for
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
 CC region and the immediate 5' tip of the coding sequence are likely to be
 CC the critical sequences for predicting PspA cross-reactions and vaccine
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 204 AA;
 Query Match 88.4%; Score 426; DB 2; Length 204;
 Best Local Similarity 87.9%; Pred. No. 5.7e-33;
 Matches 87; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
 QY 1 LKIDSDSDYAKGERAPLOSELDTKKAKLKLKLELSGKIELDAEIXEVLKDAE 60
 Db 1 IXEXDESXSEDYKLEGLRAPLOSKLDTKKAKLSKLELSGKIDELDAEIXEVLKDAE 60
 QY 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDE 99
 Db 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDE 99
 RESULT 14
 ABW02612
 ID ABW02612 standard; protein; 204 AA.
 XX
 AC ABW02612;
 XX
 DT 12-FEB-2004 (first entry)
 DE
 DE Rct123c pneumococcal surface protein A (PspA) central region.
 XX
 KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
 KW immunological; gene therapy; immunostimulant.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1.204 /note= "Xaa = Unknown amino acid"
 FT
 XX US6592876-B1.
 XX
 PD 15-JUL-2003.
 XX
 PF 15-SEP-1995; 95US-00529055.
 XX
 PR 20-APR-1993; 93US-00048896.
 PR 06-JUN-1995; 95US-00465746.
 XX
 XX (UABR-) UAB RES FOUND.
 XX
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
 XX WPI; 2003-862841/80.
 XX

PT	Immunological composition for obtaining expression products used for
PT	detecting the presence of <i>Streptococcus pneumoniae</i> or its strain,
PT	comprises at least two different full length isolated gene encoding
PT	pneumococcal surface protein A.
XX	
XX	Example 6; SEQ ID NO 58; 121pp; English.
XX	
CC	The present invention relates to an immunological composition comprising
CC	at least 2 different full length isolated genes encoding pneumococcal
CC	surface protein A (pspas) from different groups based on restriction
CC	fragment polymorphism analysis. The invention is useful for obtaining
CC	expression products by recombinant techniques to detect, determine,
CC	isolate or diagnose the presence of <i>Streptococcus pneumoniae</i> or its
CC	strain. The expression product is useful for preparing antigenic,
CC	immunological or vaccine compositions, for eliciting antibodies, an
CC	immunological response (other than or additional to antibodies) or a
CC	protective response (including antibody or other immunological response
CC	by administering compositions to a host). The invention is also useful as
CC	vaccines and in gene therapy. The present sequence is Rct123c
CC	pneumococcal surface protein A (pspa) central region. This sequence is
CC	used in the exemplification of the invention
XX	
XX	Sequence 204 AA:
XX	

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

QM protein - protein search, using sw model

Run on: June 21, 2005, 09:55:14 ; Search time 18.4867 Seconds
(without alignments)
399.760 Million cell updates/sec

Title: US-10-674-755-14

Perfect score: 482

Sequence: 1 LKIDISDSYAKGERAP.....KKAELKAEADLKKAVIDEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	480	99.6	99	4	US-09-147-875A-14
2	472	97.9	206	4	US-08-529-055-54
3	472	97.9	8931	4	US-08-714-741-32
4	458	95.0	99	2	US-08-710-749-13
5	451	93.6	99	4	US-09-147-875A-15
6	448	92.9	99	4	US-09-147-875A-16
7	440	91.3	170	4	US-08-529-055-60
8	440	91.3	181	4	US-08-529-055-42
9	440	91.3	864	4	US-08-714-741-40
10	437	90.7	99	2	US-08-710-749-14
11	437	90.7	188	4	US-08-529-055-59
12	430	89.2	99	2	US-08-710-749-17
13	429.5	89.1	100	4	US-09-147-875A-10
14	428	88.8	141	4	US-09-286-981B-2
15	428	88.8	588	4	US-08-714-741-42
16	426	88.4	99	2	US-08-710-749-15
17	426	88.4	204	4	US-08-529-055-58
18	405	84.0	1231	4	US-08-714-741-41
19	386	80.1	99	2	US-08-710-749-10
20	386	80.1	99	2	US-08-710-749-11
21	386	80.1	99	4	US-09-147-875A-11
22	386	80.1	198	4	US-08-529-055-61
23	386	80.1	204	4	US-08-529-055-51
24	386	80.1	619	1	US-08-465-746-2
25	386	80.1	619	1	US-08-214-164-2
26	386	80.1	619	2	US-08-467-852A-3
27	386	80.1	619	2	US-08-246-636-2

28	386	80.1	619	2	US-08-247-491A-3	Sequence 3, Appli
29	386	80.1	619	2	US-08-319-795-2	Sequence 2, Appli
30	386	80.1	619	2	US-08-468-985-2	Sequence 2, Appli
31	386	80.1	619	3	US-08-312-949-2	Sequence 2, Appli
32	386	80.1	648	1	US-08-072-070-2	Sequence 2, Appli
33	386	80.1	648	1	US-08-469-434-2	Sequence 2, Appli
34	386	80.1	648	1	US-08-214-222-2	Sequence 2, Appli
35	386	80.1	648	2	US-08-467-852A-2	Sequence 2, Appli
36	386	80.1	648	2	US-08-468-718-2	Sequence 2, Appli
37	386	80.1	648	2	US-08-247-491A-2	Sequence 2, Appli
38	386	80.1	695	1	US-08-446-201-3	Sequence 23, Appli
39	386	80.1	695	1	US-08-127-499A-23	Sequence 23, Appli
40	386	80.1	695	1	US-08-482-847-23	Sequence 23, Appli
41	374	77.6	288	3	US-08-312-949-4	Sequence 4, Appli
42	374	77.6	288	3	US-08-446-201-4	Sequence 4, Appli
43	372.5	77.3	100	4	US-09-147-875A-12	Sequence 12, Appli
44	365	75.7	193	4	US-08-529-055-49	Sequence 49, Appli
45	363.5	75.4	289	1	US-08-072-070-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-09-147-875A-14

; Sequence 14, Application US/09147875A

; Patent No. 6638516

; GENERAL INFORMATION:

; APPLICANT: BECKER et al.

; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS

; FILE REFERENCE: 454312-2471

; CURRENT APPLICATION NUMBER: US/09/147,875A

; CURRENT FILING DATE: 1999-05-24

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 14

; LENGTH: 99

; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: (1)..(99)

; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid

US-09-147-875A-14

Query Match 99.6%; Score 480; DB 4; Length 99;

Best Local Similarity 100.0%; Pred. No. 5.4e-40;

Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKIDISDSYAKGERAPLQSELDTKKAKLKLKLELKGKIELDAEIXELVQLKDAE 60

Db 1 LKIDISDSYAKGERAPLQSELDTKKAKLKLKLELKGKIELDAEIXELVQLKDAE 60

QY 61 GNNVAYFKEGLEKTTAKAELEKAEADLKKAVIDEPE 99

Db 61 GNNVAYFKEGLEKTTAKAELEKAEADLKKAVIDEPE 99

RESULT 2

US-08-529-055-54

; Sequence 54, Application US/08529055

; Patent No. 6592876

; GENERAL INFORMATION:

; APPLICANT: Briles, David E.

; APPLICANT: McDaniel, Larry S.

; APPLICANT: Swiatlo, Edwin

; APPLICANT: Yother, Janet

; TITLE OF INVENTION: Pneumococcal Genes, Portions

; TITLE OF INVENTION: Thereof, Expression Products

; TITLE OF INVENTION: Thereof, and Uses of Such Genes,

; TITLE OF INVENTION: Portions and Products

; NUMBER OF SEQUENCES: 73

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 206 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-54

Query Match          97.9%; Score 472; DB 4; Length 206;
Best Local Similarity 99.0%; Pred. No. 7.8e-39;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKIDSDSDYAKGGRAPLQSELDTKKAKLLKLELSGKIEELDAAIXELEVQLKDAE 60
Db 1 LKIDSDSDYAKGGRAPLQSELDTKKAKLLKLELSGKIEELDAAIXELEVQLKDAE 60

Qy 61 GNNVYAYFKGLEKTTAAEKKAELKAEADLKKAVDEPE 99
Db 61 GNNVYAYFKGLEKTTAAEKKAELKAEADLKKAVDEPE 99

RESULT 3
US-08-714-741-32
; Sequence 32, Application US/08714741
; Patent No. 650613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8991 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-714-741-32

Query Match          97.9%; Score 472; DB 4; Length 8991;
Best Local Similarity 99.0%; Pred. No. 6.6e-37;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKIDSDSDYAKGGRAPLQSELDTKKAKLLKLELSGKIEELDAAIXELEVQLKDAE 60
Db 5888 LKIDSDSDYAKGGRAPLQSELDTKKAKLLKLELSGKIEELDAAIXELEVQLKDAE 5947

Qy 61 GNNVYAYFKGLEKTTAAEKKAELKAEADLKKAVDEPE 99
Db 5948 GNNVYAYFKGLEKTTAAEKKAELKAEADLKKAVDEPE 5986

RESULT 4
US-08-710-749-13
; Sequence 13, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
```

```

; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-13

Query Match      95.0%; Score 458; DB 2; Length 99;
Best Local Similarity 96.0%; Pred. No. 7.7e-38;
Matches 95; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 LKIDSDSDYAKGERAPLQSELDTKKAKLKLKLELSGKIEELDAAIXEVLQKDAE 60
DB 1 LKIDSDSDYAKGERAPLQSELDTKKAKLKLKLELSGKIEELDAAIXEVLQKDAE 60

QY 61 GNNVYAYFKGLEKTTAAKKAELKAEADLKKAVDEPE 99
DB 61 GNNVYAYFKGLEKTTAAKKAELKAEADLKKAVDEPE 99

RESULT 5
US-09-147-875A-15
; Sequence 15, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-15

Query Match      93.6%; Score 451; DB 4; Length 99;
Best Local Similarity 92.9%; Pred. No. 3.7e-37;
Matches 92; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 LKIDSDSDYAKGERAPLQSELDTKKAKLKLKLELSGKIEELDAAIXEVLQKDAE 60
DB 1 LKIDSDSDYAKGERAPLQSELDTKKAKLKLKLELSGKIEELDAAIXEVLQKDAE 60

QY 61 GNNVYAYFKGLEKTTAAKKAELKAEADLKKAVDEPE 99
DB 61 GNNVYAYFKGLEKTTAAKKAELKAEADLKKAVDEPE 99

RESULT 6
US-09-147-875A-16
; Sequence 16, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-16

Query Match      92.9%; Score 448; DB 4; Length 99;
Best Local Similarity 91.9%; Pred. No. 7.3e-37;
Matches 91; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 LKIDSDSDYAKGERAPLQSELDTKKAKLKLKLELSGKIEELDAAIXEVLQKDAE 60
DB 1 LKIDSDSDYAKGERAPLQSELDTKKAKLKLKLELSGKIEELDAAIXEVLQKDAE 60

QY 61 GNNVYAYFKGLEKTTAAKKAELKAEADLKKAVDEPE 99
DB 61 GNNVYAYFKGLEKTTAAKKAELKAEADLKKAVDEPE 99

RESULT 7
US-08-529-055-60
; Sequence 60, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-60

Query Match      91.3%; Score 440; DB 4; Length 170;
Best Local Similarity 90.9%; Pred. No. 8.4e-36;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKIDSDSDYAKGERAPLQSELDTKKAKLKLKLELSGKIEELDAAIXEVLQKDAE 60
DB 1 LKIDSDSDYAKGERAPLQSELDTKKAKLKLKLELSGKIEELDAAIXEVLQKDAE 60

QY 61 GNNVYAYFKGLEKTTAAKKAELKAEADLKKAVDEPE 99
DB 61 GNNVYAYFKGLEKTTAAKKAELKAEADLKKAVDEPE 99

RESULT 8
US-08-529-055-42
; Sequence 42, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.

```

```
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-0712
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 181 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-42

Query Match 91.3%; Score 440; DB 4; Length 181;
Best Local Similarity 90.9%; Pred. No. 9.1e-36;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKIDSDSDYAKGGRAPLOSBLDTKAKLLKLELSGKIBELDAEIXEVLQKDAE 60
Db 1 LKEIDSDSDYKLGGRAPLQSKLDTKKAKLSKLELSKIDELDAEIXEVLQKDAE 60

Qy 61 GNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 61 GNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 9
US-08-714-741-40
; Sequence 40, Application US/08/14741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue

; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 864 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-714-741-40

Query Match 91.3%; Score 440; DB 4; Length 864;
Best Local Similarity 90.9%; Pred. No. 5.7e-35;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKIDSDSDYAKGGRAPLOSBLDTKAKLLKLELSGKIBELDAEIXEVLQKDAE 60
Db 465 LKEIDSDSDYKLGGRAPLQSKLDTKKAKLSKLELSKIDELDAEIXEVLQKDAE 524

Qy 61 GNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 525 GNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 563

RESULT 10
US-08-710-749-14
; Sequence 14, Application US/08/10749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 99 amino acids
TYPE: amino acid
STRANDEDNESS: n/a
TOPOLOGY: linear
MOLECULE TYPE: amino acid
US-08-710-749-14

Query Match 90.7%; Score 437; DB 2; Length 99;
Best Local Similarity 89.9%; Pred. No. 8.8e-36; Indels 0; Gaps 0;
Matches 89; Conservative 4; Mismatches 6;

QY 1 LKIDSDSDYAKGERAPLQSELDTKKAKLLKLELSGKIBELDAEIXEVLQKDAE 60
||:|||||
DB 1 LKIDSDSDYAKGERAPLQSKLDAKAKLLKLELSGKIBELDAEIXEVLQKXVE 60
||:|||||
QY 61 GNNVAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99
|||||
DB 61 GNNVAYFKEGLEKTTAEKATELEKAEADLKKAVDEPE 99
|||||

RESULT 11
US-08-529-055-59
; Sequence 59, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-59

Query Match 90.7%; Score 437; DB 4; Length 188;

Best Local Similarity 89.9%; Pred. No. 1.9e-35;
Matches 89; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
QY 1 LKIDSDSDYAKGERAPLQSELDTKKAKLLKLELSGKIBELDAEIXEVLQKDAE 60
||:|||||
DB 1 LKIDSDSDYAKGERAPLQSKLDTKAKLSKLELSGKIBELDAEIXEVLQKDAE 60
||:|||||
QY 61 GNNVAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99
|||||
DB 61 GNNVAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99
|||||

RESULT 12
US-08-710-749-17
; Sequence 17, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-17

Query Match 89.2%; Score 430; DB 2; Length 99;
Best Local Similarity 88.9%; Pred. No. 4.3e-35;
Matches 88; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 LKIDSDSDYAKGERAPLQSELDTKKAKLLKLELSGKIBELDAEIXEVLQKDAE 60
||:|||||
DB 1 LKIDSDSDYAKGERAPLQSKLDAKAKLSKLELSGKIBELDAEIXEVLQKDAE 60
||:|||||
QY 61 GNNVAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99
|||||
DB 61 GNNVAYFKEGLEKTTAEKATELEKAEADLKKAVDEPE 99
|||||

RESULT 13
US-09-147-875A-10
; Sequence 10, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:

```
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-147-875A-10

Query Match      89.1%; Score 429.5; DB 4; Length 100;
Best Local Similarity 90.0%; Pred. No. 4.8e-35;
Matches 90; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

Qy 1 LKIDSDSDYAKGGRAPLQSLDTTKAKLLKLELSKIEELDAEIXELE-VQKDA 59
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 LKEIDSDSDYAKGGRAPLQSLDKDAKAKLSKLELSKIDELDAEIAKLECVQKDA 60

Qy 60 EGNNVAYFKGLEKTTAKKAELEKAEADLKKAVDE 99
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 EGNNVAYFKGLEKTTAKKAELEKAEADLKKAVDE 100

RESULT 14
US-09-286-981B-2
; Sequence 2, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizemann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; PRIOR FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-286-981B-2

Query Match      88.8%; Score 428; DB 4; Length 141;
Best Local Similarity 90.7%; Pred. No. 1e-34;
Matches 88; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKIDSDSDYAKGGRAPLQSLDTTKAKLLKLELSKIEELDAEIXELEVLQKDAE 60
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 45 LKEIDSDSDYKLEGRAPLQSLDTTKAKLSKLELSKIDELDAEIAKLEVLQKDAE 104

Qy 61 GNNVAYFKGLEKTTAKKAELEKAEADLKKAVDE 97
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 105 GNNVAYFKGLEKTTAKKAELEKAEADLKKAVDE 141

RESULT 15
US-08-714-741-42
; Sequence 42, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
```

```
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 42:
; LENGTH: 588 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-714-741-42

Query Match      88.8%; Score 428; DB 4; Length 588;
Best Local Similarity 90.7%; Pred. No. 5.4e-34;
Matches 88; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKIDSDSDYAKGGRAPLQSLDTTKAKLLKLELSKIEELDAEIXELEVLQKDAE 60
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 492 LKEIDSDSDYKLEGRAPLQSLDTTKAKLSKLELSKIDELDAEIAKLEVLQKDAE 551

Qy 61 GNNVAYFKGLEKTTAKKAELEKAEADLKKAVDE 97
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 552 GNNVAYFKGLEKTTAKKAELEKAEADLKKAVDE 588

Search completed: June 21, 2005, 10:25:20
Job time : 18.4867 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 10:12:15 ; Search time 63.2388 Seconds
(without alignments)
601.118 Million cell updates/sec

Title: US-10-674-755-14

Perfect score: 482

Sequence: 1 LKDIQSDSDYAKGERAP.....KKAELKAEADLKKADEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	480	99.6	99	15	US-10-674-755-14
2	472	97.9	206	15	US-10-299-636-69
3	451	93.6	99	15	US-10-674-755-15
4	448	92.9	99	15	US-10-674-755-16
5	440	91.3	170	15	US-10-299-636-75
6	440	91.3	181	15	US-10-299-636-57
7	440	91.3	643	15	US-10-299-636-95
8	440	91.3	670	9	US-09-748-875-63
9	440	91.3	670	10	US-09-298-523B-63
10	440	91.3	690	9	US-09-748-875-61
11	440	91.3	690	10	US-09-298-523B-61

12	440	91.3	691	9	US-09-748-875-1	Sequence 1, Appli
13	440	91.3	691	10	US-09-298-523B-1	Sequence 1, Appli
14	440	91.3	701	9	US-09-748-875-62	Sequence 62, Appl
15	440	91.3	701	10	US-09-298-523B-62	Sequence 62, Appl
16	440	91.3	707	9	US-09-748-875-2	Sequence 2, Appli
17	440	91.3	707	10	US-09-298-523B-2	Sequence 2, Appli
18	440	91.3	711	9	US-09-748-875-3	Sequence 3, Appli
19	440	91.3	711	10	US-09-298-523B-3	Sequence 3, Appli
20	440	91.3	739	17	US-10-732-923-3294	Sequence 3294, Ap
21	440	91.3	929	9	US-09-748-875-60	Sequence 60, Appl
22	440	91.3	929	10	US-09-298-523B-60	Sequence 60, Appl
23	440	91.3	929	15	US-10-299-636-94	Sequence 94, Appl
24	437	90.7	188	15	US-10-299-636-74	Sequence 74, Appl
25	429.5	89.1	100	15	US-10-674-755-10	Sequence 10, Appl
26	428	88.8	141	14	US-10-254-995-2	Sequence 14, Appl
27	428	88.8	589	9	US-09-748-875-14	Sequence 14, Appl
28	428	88.8	589	10	US-09-298-523B-14	Sequence 97, Appl
29	428	88.8	589	15	US-10-299-636-97	Sequence 73, Appl
30	426	88.4	204	15	US-10-299-636-73	Sequence 11, Appl
31	386	80.1	99	15	US-10-674-755-11	Sequence 11, Appl
32	386	80.1	198	15	US-10-299-636-76	Sequence 76, Appl
33	386	80.1	204	15	US-10-299-636-66	Sequence 66, Appl
34	386	80.1	354	15	US-10-299-636-105	Sequence 105, App
35	386	80.1	588	15	US-10-299-636-96	Sequence 96, Appl
36	386	80.1	619	10	US-09-882-774-1	Sequence 1, Appli
37	386	80.1	619	15	US-10-282-122A-73702	Sequence 73702, A
38	386	80.1	619	16	US-10-414-532-72	Sequence 12, Appl
39	372.5	77.3	100	15	US-10-674-755-12	Sequence 12, Appl
40	365	75.7	193	15	US-10-299-636-64	Sequence 64, Appl
41	361	74.9	99	15	US-10-674-755-17	Sequence 17, Appl
42	359	74.5	195	15	US-10-299-636-86	Sequence 86, Appl
43	355.5	73.8	336	15	US-10-299-636-103	Sequence 103, App
44	355	73.7	99	15	US-10-674-755-13	Sequence 13, Appl
45	308.5	64.0	100	15	US-10-674-755-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-10-674-755-14
; Sequence 14, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; PRIOR FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(99)
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid
US-10-674-755-14

Query Match	99.6%	Score 480;	DB 15;	Length 99;
Best Local Similarity	100.0%	Pred. No. 3.5e-33;		
Matches	99;	Conservative	0;	Mismatches 0;
Indels	0;	Gaps	0;	
Qy	1	LKDIQSDSDYAKGERAPLQSELDTKKAKLLKLBELSGKIEELDAIXELVQLKDAE	60	
Db	1	LKDIQSDSDYAKGERAPLQSELDTKKAKLLKLBELSGKIEELDAIXELVQLKDAE	60	
Qy	61	GNNVAYPKGLEKTTAKKAELEKAEADLKKADEPE	99	

[illegible]

RESULT 6

```

US-10-299-636-57
; Sequence 57, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatcio, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-57

```

```

Query Match      91.3%; Score 440; DB 15; Length 181;
Best Local Similarity 90.9%; Pred. No. 1.6e-29;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKQIDSDSDYAKGGRAPQSQSLDTTKAKLLKLEELSGKIELDAAIXEILEVQLKDAE 60
    |||||
Db 1 LKEIDSDSDYLKEGRLAPQSQSLDTTKAKLLKLEELSDKIDELDAIAKLEVQLKDAE 60
    |||||

QY 61 GNNNVYAFKEGLEKTTAAEKKAELEKAEADLKKAVDPE 99
    |||||
Db 61 GNNNVYAFKEGLEKTTAAEKKAELEKAEADLKKAVDPE 99
    |||||

```

RESULT 7

```

US/10-299-636-95
; Sequence 95, Application US/10299636
; Publication No. US2004007847A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
;
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
;
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055

```

```

; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatenIn Ver. 2.1
; SEQ ID NO 95
; LENGTH: 643
; TYPE: PRN
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-95

```

Query Match 91.3%; Score 440; DB 15; Length 643;
Best Local Similarity 90.9%; Pred. No. 6.7e-29;
Matches 90: Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy	1	LKDIDSDSDYAKGERAPLOSELDTKAKLLKEELSGKIELDIAEIXELEVLQDAE	60
		::::: :	
Db	245	LKEIDESDSDDYLKEGLRPLAPLQSLDTKAKUKSLEBLSKDIDELDIAEIAKLEIVQLQDAE	304
		::::: :	
Qy	61	GNNNVAYPKFEGLEKTTAAKKAELEKAADLKAVDEP	99
		::::: :	
Db	305	GNNNVAYPKFEGLEKTTAAKKAELEKAADLKAVDEP	343
		::::: :	

RESULT 8

```

US-09-748-875-63
; Sequence 63, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748, 875
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 670
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-63

```

```
Query Match          91.3%; Score 440; DB 9; Length 670;
Best Local Similarity 90.9%; Pred.No. 7.1e-23;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
```

```
Qy      1 LKIDSDSEDYAKGERAPLQSELTTKYAKLLKLSELGSKIIBELDAEIXELFVQLKDAE 60
|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db     498 LKEIDSDSEDYLKEGLRAPLQSKLDTTKYAKLSKLEELSDKIDELDAEIAKLEFVQLKDAE 557
```

```
Qy      61 GNNNVZAYFKEGLEKTTAEKKAELEKABADLKKAVIDEPE 99
|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db     558 GNNNVZAYFKEGLEKTTAEKKAELEKABADLKKAVIDEPE 596
```

PRECISE

```

RESUL 3
US-09-298-523B-63
; Sequence 63, Application US/09298523B
; Publication No. US20030059438A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC),
; EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/298,523B
; CURRENT FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 670
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae

```


RESULT 14
US-09-748-875-62
; Sequence 62, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: US/09/748,875
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 701
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-62

Query Match 91.3%; Score 440; DB 9; Length 701;
Best Local Similarity 90.9%; Pred. No. 7.4e-29;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 1 LKIDSDSDYAKGGERAPLQSELDTKKAKLLKLELSGKIEELDAAIXELEVQLKDAE 60
DB 529 LKEIDSDSDYAKGGERAPLQSELDTKKAKLLKLELSGKIEELDAAIXELEVQLKDAE 588
QY 61 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 99
DB 589 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 627

RESULT 15
US-09-298-523B-62
; Sequence 62, Application US/09298523B
; Publication No. US20030059438A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 701
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-62

Query Match 91.3%; Score 440; DB 10; Length 701;
Best Local Similarity 90.9%; Pred. No. 7.4e-29;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 1 LKIDSDSDYAKGGERAPLQSELDTKKAKLLKLELSGKIEELDAAIXELEVQLKDAE 60
DB 529 LKEIDSDSDYAKGGERAPLQSELDTKKAKLLKLELSGKIEELDAAIXELEVQLKDAE 588
QY 61 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 99
DB 589 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 627

Search completed: June 21, 2005, 11:18:35
Job time : 64.2388 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:53:39 ; Search time 9.9 Seconds
(without alignments)
962.168 Million cell updates/sec

Title: US-10-674-755-14

Perfect score: 482

Sequence: 1 LKXIDSDSDYAKGERAP.....KKALEKADLKKAVDEPE 99

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	386	80.1	619	2 A97887	surface protein ps
2	386	80.1	619	2 A41971	surface protein ps
3	119.5	24.8	161	2 S48396	tropomyosin TPM2 -
4	116.5	24.2	744	2 P95013	pneumococcal surfa
5	112.5	23.3	281	2 F75216	hypothetical prote
6	108.5	22.5	650	2 A11333	ABC transporter (A
7	108.5	22.5	1006	2 C70445	ATPase subunit of
8	107	22.2	1319	2 A28313	glued protein - fr
9	104.5	21.7	852	2 D72230	conserved hypothet
10	103.5	21.5	785	2 T01025	hypothetical prote
11	103.5	21.5	1110	2 I51116	NF-180 - sea lamp
12	102	21.2	764	2 T05409	hypothetical prote
13	101.5	21.1	279	2 D71453	hypothetical prote
14	100	20.7	1169	2 D71453	P115 homolog - Met
15	100	20.7	1269	2 F84730	probable myosin he
16	99	20.5	233	2 S70531	bbk2.11 protein pr
17	99	20.5	650	2 A41704	ABC transporter (A
18	99	20.5	1177	2 T05150	chromosome segrega
19	98.5	20.4	779	2 T05990	hypothetical prote
20	97	20.1	1790	2 S67593	transport protein
21	97	20.1	2139	2 T18296	myosin heavy chain
22	96.5	20.0	384	2 G86287	hypothetical prote
23	96	19.9	1410	1 A57013	early endosome ant
24	96	19.9	1875	2 S38173	myosin-like protei
25	96	19.9	3488	2 T34418	hypothetical prote
26	95.5	19.8	229	2 S70532	outer surface prot
27	95.5	19.8	1164	2 T24806	hypothetical prote
28	95.5	19.8	1179	2 F21190	probable chromosom
29	95	19.7	876	2 A23767	myosin heavy chain

30	95	19.7	880	2 F75103	conserved hypothet
31	95	19.7	1078	2 T18352	protein P120 - Myc
32	95	19.7	1475	2 T33318	hypothetical prote
33	95	19.7	1690	2 T13030	microtubule bindin
34	95	19.7	1976	2 A59252	myosin heavy chain
35	94.5	19.6	392	2 G95258	secreted 45 kd pro
36	94.5	19.6	392	2 B98124	general stress pro
37	94.5	19.6	488	2 F97039	hypothetical prote
38	94.5	19.6	635	2 A10625	ABC transporter AT
39	94.5	19.6	1170	2 A56157	chromosome segrega
40	94.5	19.6	1805	1 A64224	hypothetical prote
41	94	19.5	407	1 EDBBQ3	immediate-early pr
42	94	19.5	1156	2 E69444	chromosome segrega
43	94	19.5	1539	2 T18372	repeat organellar
44	93.5	19.4	629	2 F86351	protein T26F17.2 [
45	93.5	19.4	886	2 H69378	conserved hypothet

ALIGNMENTS

RESULT 1

A97887
surface protein pspA precursor [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: A97887
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaekunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: A97887
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-619 <KUR>
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:AE007317; PIDN:AAK98925.1; PID:glf
C:Genetics:
A:Gene: pspA

Query Match	80.1%	Score	386;	DB	2;	Length	619;
Best Local Similarity	81.8%	Pred. No.	1.2e-19;				
Matches	81;	Conservative	6;	Mismatches	12;	Indels	0;
		Gaps	0;				
QY	1	LKXIDSDSDYAKGERAPLOSLDTKKALKLEELSGKTEELDAEIXELEVQLKDAE	60				
DB	223	LKEIDSESDYAKGFRAPLQSLDKAKKLEELSDKIDELDAEIAKLEQDKAAE	282				
QY	61	GNNVVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE	99				
DB	283	ENNVVEDYFKEGLEKTTIAKAELEKAEADLKKAVNEPE	321				

RESULT 2

A41971
surface protein pspA precursor - Streptococcus pneumoniae
N:Alternate names: pneumococcal surface protein A
C:Species: Streptococcus pneumoniae
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A41971; A60282; A33134
R:Yother, J.; Briles, D.E.
J. Bacteriol. 174, 601-609, 1992
A:Title: Structural properties and evolutionary relationships of PspA, a surface protein
A:Reference number: A41971; MUID:92105030; PMID:1729249
A:Accession: A41971
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-619 <VOT>
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:M74122; NID:g153840; PIDN:AAA27018
A>Note: sequence extracted from NCBI backbone (NCBI:75635, NCBI:P:75636)
R:Talkington, D.F.; Crimmins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.

A11333
ABC transporter (ATP-binding protein) homolog lmo2073 [imported] - Listeria monocytogenes
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: A11333
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Pshi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Makok, C.; Schluteter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
A:Reference number: A1077; MUID:21537279; PMID:11679669
A:Accession: A11333
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-650 <GLA>
A:Cross-references: UNIPROT:Q8Y519; GB:NC_003210; PIDN:CAD00151.1; PID:g16411543; GSPDB:
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo2073
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
Query Match 22.5%; Score 108.5; DB 2; Length 650;
Best Local Similarity 32.7%; Pred. No. 1.8; Mismatches 21; Indels 21; Gaps 4;
Matches 36; Conservative 21; Mismatches 32; Indels 32; Gaps 4;
QY 2 KQIDSDSEDYAKGGERA-----PLQSELDTKKAKLL-----KLEELSGKIEELDAEI 49
Db 541 KELLARLDAEDRRKKGQEVATASVRKLNQEEKEQKLLRQRKKLEIEKSNWEETDEKI 600
QY 50 XELEVQLKDAEGNNVVEAFYFKEGLEKTTAEKKAELKAEADLKKAVIDE 99
Db 601 AELEQLTNPE-----VFQDHEKALEIT-----QELDAVRADGKELMEWE 641
RESULT 7
C70445
ATPase subunit of ATP-dependent proteinase (EC 3.4.-.-) - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C:Accession: C70445
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98156666; PMID:9537320
A:Accession: C70445
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1006 <AQF>
A:Cross-references: UNIPROT:O67589; GB:AE000750; NID:g2983999; PIDN:AAC07550.1; PID:g298
A:Experimental source: strain VF5
C:Genetics:
A:Gene: ctpB
C:Superfamily: endopeptidase Clp ATP-binding chain
C:Keywords: hydrolase
Query Match 22.5%; Score 108.5; DB 2; Length 1006;
Best Local Similarity 33.6%; Pred. No. 2.8; Mismatches 27; Indels 23; Gaps 5;
Matches 36; Conservative 21; Mismatches 27; Indels 23; Gaps 5;
QY 1 LKIDISDSSE-----DYAKGERAPLQSELDTKKAKLLK-----LEELSGKIEELDAEI 54
Db 552 IKALEEQIIEANLKGDEKE-----AQLKIEKAKLEKEQBLGKVGVEAKIAELKK 604
QY 55 QLKDAEGNNVVEAFYFKEGLEKTTAEKKAELKAEADLKKAVIDE 97
Db 605 KIEE-----LDEKIEAEKGDVEAEALIEKAKLEKELKKLEQ 645
RESULT 8
A28313
glued protein - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C:Accession: A28313
R:Swaroop, A.; Swaroop, M.; Garen, A.
Proc. Natl. Acad. Sci. U.S.A. 84, 6501-6505, 1987
A:Title: Sequence analysis of the complete cDNA and encoded polypeptide for the glued ger
A:Reference number: A28313; MUID:87317680; PMID:2819881
A:Accession: A28313
A:Molecule type: DNA; mRNA
A:Residues: 1-1319 <SWA>
A:Cross-references: UNIPROT:PI3496
A:Note: the authors' translation is inconsistent with the nucleotide sequence in the regi
C:Genetics:
A:Gene: FlyBase:Gl
A:Cross-references: FlyBase:FBgn0001108
A:Introns: 18/2; 479/3
C:Keywords: cytoskeleton; glycoprotein
F:397,590,771,888,980,1110,1127,1133,1142/Binding site: carboxydrate (Asn) (covalent) #st
Query Match 22.2%; Score 107; DB 2; Length 1319;
Best Local Similarity 32.7%; Pred. No. 4.6; Mismatches 19; Indels 22; Gaps 4;
Matches 33; Conservative 19; Mismatches 27; Indels 22; Gaps 4;
QY 1 LKIDISDSSEDYAKGERAPLQSELDTKKAKLLK-----EELSGKIEELDAEI 56
Db 429 LRDLSAHDKHDYQK-----LSKELEMKRSEVTELETKELSAKIDELEIAIVADLQEQV 482
QY 57 KDAEGNNVVEAFYFKEGLEKTTAEKKAELKAEADLKKAVIDE 97
Db 483 DAALG-----AEEMVEQLAKKQWLE-----DKVKLEE 511
RESULT 9
D72230
conserved hypothetical protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: D72230
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, G.
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: D72230
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-852 <ARN>
A:Cross-references: UNIPROT:Q9X1X1; GB:AE001806; GB:AE000512; NID:g4982196; PIDN:AAD3670:
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM1636
C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032
Query Match 21.7%; Score 104.5; DB 2; Length 852;
Best Local Similarity 30.3%; Pred. No. 4.4; Mismatches 20; Indels 11; Gaps 2;
Matches 27; Conservative 20; Mismatches 31; Indels 11; Gaps 2;
QY 6 ESDSEDYAKGERAPLQSELDTKKAKLLKLE-----LSGKIEELDAEI 61
Db 506 EKIEELHRLGYSDELLEKLDKRRKKLIEERHSISQKITAAADVQISQIENQLKEIKG 565
QY 62 NNVEAFYFKEGLEKTTAEKKAELKAEAD 90
Db 566 E-----IEAKRETUKEQREMDQLKSD 587
RESULT 10
T01025
hypothetical protein YUP8H12R.8 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004

C;Accession: T01025
R;Theologian, A.; Vysotskaia, V.S.; Osborne, B.I.; Schwartz, J.R.; Federspiel, N.A.; Kwan
Oefner, P.; Davis, R.W.
submitted to the EMBL Data Library, May 1998
A;Description: Arabidopsis thaliana chromosome 1 YAC YUP8H12R sequence.
A;Reference number: Z14227
A;Accession: T01025
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-785 <THE>
A;Cross-references: UNIPROT:O64521; EMBL:AC002986; NID:g2494106; PID:g3152577; GSPDB:GNQ
A;Experimental source: cultivar Columbia
C;Genetics:
A;Gene: ATSP:YUP8H12R.8
A;Map position: 1
A;Introns: 108/3; 157/3; 190/3; 341/3; 394/3; 420/3; 455/2; 495/3; 526/3; 543/1; 577/3;
Query Match 21.5%; Score 103.5; DB 2; Length 785;
Best Local Similarity 26.5%; Pred. No. 4.8;
Matches 31; Conservative 27; Mismatches 28; Indels 31; Gaps 5;
QY 9 SEDYAK-EGBRAPLQSLDTKKAKLLKLE-----ELSGKIEELDAEIXE 51
Db 65 SQPFSSLESQNAKLQSFDDRLAEAGSQAKQHLQSVPCVIEKDGVEVMSTEMSE 124
QY 52 LE-----VOLKDA---EGNNNVNPFKE--GLEKTTTAAKKAELKAEADLKA 94
Db 125 LHKSKRQLMELLEQKDAEISEKSNSTIKSYLDKIVKLTDTSSEKARLAEATAELARS 181
RESULT 11
I51116
NF-180 - sea lamprey
C;Species: Petromyzon marinus (sea lamprey)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
R;Accession: I51116
R;Accession: A.J.; Kamholz, J.; Selzer, M.E.
Brain Res. Mol. Brain Res. 29, 43-52, 1995
A;Title: The single lamprey neurofilament subunit (NF-180) lacks multiphosphorylation re
A;Reference number: I51116; MUID:95287814; PMID:7770000
A;Accession: I51116
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1110 <JAC>
A;Cross-references: UNIPROT:Q91255; EMBL:U19361; NID:g632548; PIDN:AAA80106.1; PID:g6325
C;Superfamily: neurofilament triplet H protein
Query Match 21.5%; Score 103.5; DB 2; Length 1110;
Best Local Similarity 35.0%; Pred. No. 6.7;
Matches 35; Conservative 19; Mismatches 39; Indels 7; Gaps 5;
QY 3 DIDESSEDYA-KEGE-RAPLQSELDTKAKLLKLELSGKIEELDAEIXEVLKDAE 60
Db 502 EABEEEDBGRKEGEAAEAEAEAEVEEKEAEAEAEAEAEAEAEAEAEAEAEAE 558
QY 61 GNNNVNPFKEGLEKTTTAAKKAELKAE-ADLKKAVDPE 99
Db 559 GESEAEAEGEAEAEAEVEEAE-LEKAEAEAEAEVEEEE 597
RESULT 12
T05409
hypothetical protein F10M6.170 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T05409
R;Bavan, M.; Weichselgartner, M.; Fattmann, B.; Granderrath, K.; Dauner, D.; Herzl, A.; N
submitted to the Protein Sequence Database, February 1998
A;Reference number: Z15414
A;Accession: T05409
A;Molecule type: DNA
A;Residues: 1-764 <BEV>
A;Cross-references: UNIPROT:O49371; EMBL:AL021811

Query Match 20.7%; Score 100; DB 2; Length 1169;
Best Local Similarity 32.4%; Pred. No. 12;
Matches 33; Conservative 22; Mismatches 39; Indels 8; Gaps 3;
QY 1 LKIDIDESDYAKGERAPLQSELDTKKAKLKLEELSGKIEELDAEIXEVLQKDAE 60
DB 799 LKRMNIEGELKILEKEKAKKNEID--KGLTLVKELIPKIEELNKKVSELINKKVILE 856
QY 61 GNNVVEAYPKGLEKTTA---EKKAELKAEADLKAVDEPE 99
DB 857 KN---ISFYKESIEKNLSILEEKRYEELAKNKLKELTEKKE 895

RESULT 15
F84730
A:probable myosin heavy chain [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: F84730
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: F84730
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1269 <STO>
A:Cross-references: GB:AE002093; NID:G6598483; PIDN:AAC69932.2; GSPDB:GN00139
C:Genetics:
A:Gene: At2g32240
A:Map position: 2

Query Match 20.7%; Score 100; DB 2; Length 1269;
Best Local Similarity 26.3%; Pred. No. 13;
Matches 36; Conservative 24; Mismatches 39; Indels 38; Gaps 4;
QY 1 LKIDIDESDYAK-----GERAPLQSELDTKKAKLKLE-----EL 38
DB 263 IKELNEKMSNEKVEALKSSAGELAAVQELALSKSRLLTEQKVSSTEALIDELTQEL 322
QY 39 SGK-----IEELDAEIXEVLQKDAEAGNN---VEAYFKGLEKTTAEKKA 82
DB 323 EQKKASESRPKBELSVLQDLDAQTGKLQAKLSEQEGINSKLAEELKEKELLESLSKDQBE 382
QY 83 ELEKAEADLKAVDEPE 99
DB 383 KLTANEKLAEVLKEKE 399

Search completed: June 21, 2005, 10:12:00
Job time : 10.9 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:54:24 ; Search time 61.3194 Seconds
(without alignments)
826.751 Million cell updates/sec

Title: US-10-674-755-14

Perfect score: 482

Sequence: 1 LKDIQSDSEDYAKGERAP.....KKAELEKADLKKAVDEPE 99

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	464	96.3	224	2 Q8GNS8	Q8GNS8 streptococc
2	464	96.3	249	2 Q9L575	Q9L575 streptococc
3	464	96.3	426	2 Q9LAY5	Q9LAY5 streptococc
4	460	95.4	99	2 Q8KQK4	Q8KQK4 streptococc
5	451	93.6	395	2 Q9LAY2	Q9LAY2 streptococc
6	451	93.6	408	2 Q9LAY0	Q9LAY0 streptococc
7	440	91.3	739	2 Q9RQT4	Q9RQT4 streptococc
8	440	91.3	820	2 Q9RQT1	Q9RQT1 streptococc
9	440	91.3	929	2 Q9KK19	Q9KK19 streptococc
10	440	91.3	929	2 Q9ZAV5	Q9ZAV5 streptococc
11	427	88.6	437	2 Q9LAY4	Q9LAY4 streptococc
12	403.5	83.7	869	2 Q9KK27	Q9KK27 streptococc
13	386	80.1	417	2 Q9LAY3	Q9LAY3 streptococc
14	386	80.1	619	2 Q54972	Q54972 streptococc
15	386	80.1	619	2 Q8DR10	Q8DR10 streptococc
16	364	75.5	415	2 Q9LAY1	Q9LAY1 streptococc
17	306.5	63.6	333	2 Q9LAZ3	Q9LAZ3 streptococc
18	305.5	63.4	222	2 Q9L577	Q9L577 streptococc
19	305.5	63.4	225	2 Q9L591	Q9L591 streptococc
20	305.5	63.4	262	2 Q9L576	Q9L576 streptococc
21	305.5	63.4	415	2 Q9LAY7	Q9LAY7 streptococc
22	304.5	63.2	394	2 Q9LAY6	Q9LAY6 streptococc
23	304.5	63.2	395	2 Q9LAZ1	Q9LAZ1 streptococc
24	302.5	62.8	246	2 Q9L578	Q9L578 streptococc
25	301.5	62.6	255	2 Q9L581	Q9L581 streptococc
26	301.5	62.6	255	2 Q9L586	Q9L586 streptococc
27	298.5	61.9	416	2 Q9LAY8	Q9LAY8 streptococc
28	298.5	60.7	194	2 Q9L585	Q9L585 streptococc
29	292.5	60.7	218	2 Q6UBB2	Q6UBB2 streptococc
30	292.5	60.7	233	2 Q9L568	Q9L568 streptococc
31	292.5	60.7	236	2 Q9L569	Q9L569 streptococc

```

32 292.5 60.7 243 2 Q9L564 streptococc
33 292.5 60.7 243 2 Q9L567 streptococc
34 292.5 60.7 244 2 Q9L565 streptococc
35 292.5 60.7 247 2 Q9L566 streptococc
36 292.5 60.7 249 2 Q9L570 streptococc
37 292.5 60.7 254 2 Q9L563 streptococc
38 292.5 60.7 401 2 Q9LAZ2 streptococc
39 291.5 60.5 406 2 Q9LAZ0 streptococc
40 290.5 60.3 340 2 Q8KQK5 streptococc
41 289.5 60.1 237 2 Q9L592 streptococc
42 289.5 60.1 395 2 Q9LAY9 streptococc
43 280.5 58.2 207 2 Q8GNS9 streptococc
44 189.5 39.3 653 2 Q34097 streptococc
45 173 35.9 256 2 Q9L595 streptococc

```

ALIGNMENTS

```

RESULT 1
Q8GNS8
ID Q8GNS8 PRELIMINARY; PRT; 224 AA.
AC Q8GNS8;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PN124;
RX MEDLINE=22411996; PubMed=12354862;
RA Dicuonzo G., Gherardi G., Gertz R.E., D'Ambrosio F., Goglio A.,
RA Lorino G., Recchia S., Pantosti A., Beall B.;
RT "Genotypes of invasive pneumococcal isolates recently recovered from
RT Italian patients.";
RL J. Clin. Microbiol. 40:3660-3665(2002).
DR EMBL; AP490267; AAN37735.1; -.
DR HSSP; P00192; IAPC.
DR InterPro; IPR009082; His_kin_homodim.
FT NON_TER 1
FT NON_TER 224 224
SQ SEQUENCE 224 AA; 23418 MW; 48674E27AFB66A95 CRC64;

Query Match 96.3%; Score 464; DB 2; Length 224;
Best Local Similarity 96.0%; Pred. No. 8.6e-24;
Matches 95; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LKDIQSDSEDYAKGERAPLOSELDTKKAKLLKLEELSGKIEELDAETKELEVLKDAE 60
   |||||
DB 17 LKDIQSDSEDYVKGFRAPLOSELDTKKAKLLKLEELSGKIEELDAETKELEVLKDAE 76
   |||||
QY 61 GNNVYAEYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99
   |||||
DB 77 GNNVYAEYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 115
   |||||

RESULT 2
Q9L575
ID Q9L575 PRELIMINARY; PRT; 249 AA.
AC Q9L575;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.

```



```

    taps 0;

```

```

QY 1 LKIDIDESDYAKGERAPLQSELDTKKAKLKLKLELSGKIBELDIAIXEVLVKDAE 60
DB 530 LKIDIDESDYAKGERAPLQSELDTKKAKLKLKLELSGKIBELDIAIXEVLVKDAE 589
QY 61 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 99
DB 590 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 628

RESULT 9
Q9KK19 ID Q9KK19 PRELIMINARY; PRT; 929 AA.
AC Q9KK19
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein pspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ef10;
RX MEDLINE=21888621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae.";
RL Gene 284:63-71(2002).
DR EMBL; AF154037; AAF73809.1; -.
DR HSP; P06653; IH8G.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR Pfam; PF01473; CW binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
SQ SEQUENCE 929 AA; 105003 MW; 2DC82933032FAFA64 CRC64;

Query Match 91.3%; Score 440; DB 2; Length 929;
Best Local Similarity 90.9%; Pred. No. 1.3e-21;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKIDIDESDYAKGERAPLQSELDTKKAKLKLKLELSGKIBELDIAIXEVLVKDAE 60
DB 530 LKIDIDESDYAKGERAPLQSELDTKKAKLKLKLELSGKIBELDIAIXEVLVKDAE 589
QY 61 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 99
DB 590 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 628

RESULT 10
Q9ZAY5 ID Q9ZAY5 PRELIMINARY; PRT; 929 AA.
AC Q9ZAY5;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein C.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF6796;
RX MEDLINE=20038319; PubMed=10569772;

```

```

RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, pspC, which elicits cross-reactive antibodies to pspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; U72655; AAD00184.1; -.
DR HSP; P06653; IH8G.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR Pfam; PF01473; CW binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
SQ SEQUENCE 929 AA; 104991 MW; 2DC82933032FFB081 CRC64;

Query Match 91.3%; Score 440; DB 2; Length 929;
Best Local Similarity 90.9%; Pred. No. 1.3e-21;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKIDIDESDYAKGERAPLQSELDTKKAKLKLKLELSGKIBELDIAIXEVLVKDAE 60
DB 530 LKIDIDESDYAKGERAPLQSELDTKKAKLKLKLELSGKIBELDIAIXEVLVKDAE 589
QY 61 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 99
DB 590 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 628

RESULT 11
Q9LAY4 ID Q9LAY4 PRELIMINARY; PRT; 437 AA.
AC Q9LAY4
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL34;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071811; AAF27707.1; -.
DR InterPro; IPR002479; CW binding.
DR Pfam; PF01473; CW binding_1; 1.
DR NON TER 437
SQ SEQUENCE 437 AA; 48071 MW; F6EFD2CD13E08CD8 CRC64;

Query Match 88.6%; Score 427; DB 2; Length 437;
Best Local Similarity 88.9%; Pred. No. 4.9e-21;
Matches 88; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 LKIDIDESDYAKGERAPLQSELDTKKAKLKLKLELSGKIBELDIAIXEVLVKDAE 60
DB 235 LKIDIDESDYAKGERAPLQSELDTKKAKLKLKLELSGKIBELDIAIXEVLVKDAE 294
QY 61 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 99
DB 295 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 333

RESULT 12

```

```
Q9KK27
ID Q9KK27 PRELIMINARY; PRT; 869 AA.
AC Q9KK27;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Surface protein PspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=95;
RX MEDLINE=21888621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae.";
RL Gene 284:63-71(2002).
DR EMBL; AF154032; AAF73801.1; -.
DR HSSP; P06653; 1H8G.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpob_Ysirk.
DR InterPro; IPR007756; Rich.
DR Pfam; PF01473; CW_binding_1; 8.
DR Pfam; PF05062; Ysirk; 2.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRfams; TIGR01168; Ysirk signal; 1.
DR SEQUENCE 869 AA; 98732 MW; AFP2B504347E0220 CRC64;
SQ
Query Match 83.7%; Score 403.5; DB 2; Length 869;
Best Local Similarity 85.9%; Pred. No. 3.5e-19;
Matches 85; Conservative 3; Mismatches 10; Indels 1; Gaps 1;
QY 1 LKDIQSDSDYAKGERAPLQSELDTKKAKLLKLEELSGKIEELDAEIXEVLQKDAE 60
DB 537 LKEIDSESDYAKGERAPLQSELDTKKAKLLKLEELSGKIEELDAEIXEVLQKDAE 595
QY 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
DB 596 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 634
RESULT 13
ID Q9LAV3 PRELIMINARY; PRT; 417 AA.
AC Q9LAV3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PepA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BF10197;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PepA; mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071812; AAF27708.1; -.
DR HSSP; P00192; 256B.
DR NON TER 417
FT SEQUENCE 417 AA; 46960 MW; 876EAD3276506BEC CRC64;
SQ
Query Match 80.1%; Score 386; DB 2; Length 417;
Matches 81; Conservative 6; Mismatches 12; Indels 0; Gaps 0;
QY 1 LKDIQSDSDYAKGERAPLQSELDTKKAKLLKLEELSGKIEELDAEIXEVLQKDAE 60
DB 223 LKEIDSESDYAKGERAPLQSELDTKKAKLLKLEELSGKIEELDAEIXEVLQKDAE 282
QY 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
DB 283 ENNVYDYFKEGLEKTTAAKKAELKAEADLKKAVNEPE 321
RESULT 14
ID Q54972 PRELIMINARY; PRT; 619 AA.
AC Q54972;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pneumococcal surface protein A precursor.
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92105030; PubMed=1729249;
RA Yother J., Briles D.E.;
RT "Structural properties and evolutionary relationships of PepA, a
RT surface protein of Streptococcus pneumoniae, as revealed by sequence
RT analysis.";
RL J. Bacteriol. 174:601-609(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Yother J., Briles D.E.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; W74122; AAA27018.1; -.
DR PIR; A41971; A41971.
DR PIR; A97887; A97887.
DR HSSP; P06653; 1HCX.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam; PF01473; CW_binding_1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_3.
KW Signal.
FT SIGNAL 1 31 Potential.
FT CHAIN 32 619 pneumococcal surface protein A.
SQ SEQUENCE 619 AA; 68605 MW; 5AA8BDB40C2841CA CRC64;
Query Match 80.1%; Score 386; DB 2; Length 619;
Best Local Similarity 81.8%; Pred. No. 3.8e-18;
Matches 81; Conservative 6; Mismatches 12; Indels 0; Gaps 0;
QY 1 LKDIQSDSDYAKGERAPLQSELDTKKAKLLKLEELSGKIEELDAEIXEVLQKDAE 60
DB 223 LKEIDSESDYAKGERAPLQSELDTKKAKLLKLEELSGKIEELDAEIXEVLQKDAE 282
QY 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
DB 283 ENNVYDYFKEGLEKTTAAKKAELKAEADLKKAVNEPE 321
RESULT 15
ID Q8DR10 PRELIMINARY; PRT; 619 AA.
AC Q8DR10;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein pspA.
GN Name=pspA; Ordered locus names=spr0121;
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
```

```

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=171101;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21429245; PubMed=11544234;
RX DOI=10.1128/JB.183.19.5709-5717.2001;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAhren S.M., McHenry M., McLeaster K., Mundy C.W., Nicas T.I.,
RA Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rockey P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rosteck P.R. Jr., Skatrud P.L.,
RA Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717(2001).
DR EMBL; AE008396; AAK98925.1; -.
DR PIR; A41971; A41971.
DR PIR; A97887; A97887.
DR HSP; P06653; IHCX.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam; PF01473; CW binding_1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_3.
KW Complete proteome.
SQ SEQUENCE 619 AA; 68605 MW; 5AA8BDB40C2841CA CRC64;

Query Match      80.1%; Score 386; DB 2; Length 619;
Best Local Similarity 81.8%; Pred. No. 3.8e-18;
Matches 81; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

Qy      1 LKIDESDSDYAKEGERAPIQSELDTKAKLLKLELSCKIEELDAAIYEVLKDAE 60
Db      223 LKEIDSESDYAKEGFRAPLQSKLDKAKLCKLELSCKIDELDAEIAKLEDLKAAE 282

Qy      61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKXAVDEPE 99
Db      283 ENNVEDYFKEGLEKTTAAKKAELKTEADLKXAVNEPE 321

```

Search completed: June 21, 2005, 10:22:11
 Job time : 61.3194 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:50:39 ; Search time 73.8459 Seconds
(without alignments)
518.502 Million cell updates/sec

Title: US-10-674-755-15

Perfect score: 484

Sequence: 1 LINEINSEDSYAKGFRAP.....KKAELKARADLKVADEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : 1: Genesecp16Dec04:*

2: Genesecp1980s:*

3: Genesecp1990s:*

4: Genesecp2000s:*

5: Genesecp2001s:*

6: Genesecp2002s:*

7: Genesecp2003as:*

8: Genesecp2003bs:*

9: Genesecp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	454	93.8	206	2	AAW14574 Streptoco
2	454	93.8	206	7	ABW02608 Dbl5c pne
3	454	93.8	8991	6	ABU08487 S. pneumo
4	440	90.9	170	7	ABW02614 Rct135c p
5	440	90.9	181	7	ABW02596 0922134c
6	440	90.9	865	6	ABU08489 S. pneumo
7	440	90.9	929	2	AAW14593 Streptoco
8	440	90.9	929	2	AAW14593 Streptoco
9	437	90.3	188	7	ABW02613 Streptoco
10	437	90.3	188	7	ABW02613 Streptoco
11	429	88.6	1231	6	ABU08490 Fragment
12	428	88.4	588	6	ABU08491 Coiled co
13	428	88.4	589	2	AAW143392 PspC alph
14	426	88.0	204	2	AAW14578 Streptoco
15	426	88.0	204	7	ABW02612 Rct123c p
16	425.5	87.9	180	2	AAW14562 Streptoco
17	422.5	87.3	187	2	AAW14579 Streptoco
18	404	83.5	198	2	AAW14581 Streptoco
19	401	82.9	198	7	ABW02615 Rxic pneu
20	401	82.9	315	2	AAW043375 Streptoco
21	401	82.9	619	2	AAR63437 Pneumococ
22	401	82.9	619	2	AAR63437 Pneumococ
23	401	82.9	619	2	AAR63437 Pneumococ
24	401	82.9	619	2	AAR63437 Pneumococ
25	401	82.9	619	5	AAE18782 S. pneumo

26	401	82.9	619	6	ABU45778 Protein e
27	401	82.9	619	8	ADO52126 Streptoco
28	401	82.9	648	2	AAW70336 Streptoco
29	401	82.9	648	2	AAW62274 Streptoco
30	401	82.9	648	2	AAW41837 Streptoco
31	401	82.9	648	2	AAW87879 A pneumoc
32	401	82.9	653	2	AAW92456 S. pneumo
33	401	82.9	684	2	AAW73912 Streptoco
34	393	81.2	204	2	AAW14571 Streptoco
35	393	81.2	204	7	ABW02605 Bf1019c p
36	381	78.7	653	2	AAW27150 PspA frag
37	378.5	78.2	289	2	AAW62276 Streptoco
38	378.5	78.2	289	2	AAW41840 Streptoco
39	378.5	78.2	289	2	AAW87910 Protein s
40	378.5	78.2	289	2	AAW92458 S. pneumo
41	377	77.9	190	2	AAW14569 Streptoco
42	377	77.9	193	7	ABW02603 Bg9163c p
43	374	77.3	195	2	AAW14591 Streptoco
44	374	77.3	195	7	ABW02625 Wu2c pneu
45	357	73.8	623	6	ABU08494 Fragment

ALIGNMENTS

RESULT 1
AAW14574
ID AAW14574 standard; protein; 206 AA.
XX
AC AAW14574;
XX
DT 17-OCT-2003 (revised)
DT 28-OCT-1997 (first entry)
XX
DE Streptococcus pneumoniae PspA central region.

PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
bacteraemia; pneumonia.

Streptococcus pneumoniae; strain Dbl5.

Key Location/Qualifiers

FT Misc-difference 50 /note= "unidentified amino acid"

XX WO9709994-A1.

PN 20-MAR-1997.

PD 16-SEP-1996; 96WO-US014819.

XX 15-SEP-1995; 95US-00529055.

PR (UABR-) UAB RES FOUND.

XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;

PI Hollingshead S, Tart R, Brooks-Walter A;

DR WPI; 1997-202002/18.

XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
in vaccines for protecting animals against S.pneumoniae infection.

PS Example 6; Fig 13; 296pp; English.

CC This sequence shows the central portion, including the C-terminus of the
alpha-helix region and some of the proline-rich region, of pneumococcal
surface protein A (PspA) of Streptococcus pneumoniae strain Dbl5.

CC Comparison of the N-terminal and central regions (AAW14533-57 and
AAW14562-91) of PspA polypeptides from different pneumococcal strains can
be used to divide the strains into several families based on sequence
homologies. PspA polypeptides, or fragments of them, can be used in
vaccines to protect animals against S. pneumoniae infection and hence for

CC the prevention of diseases such as otitis media, meningitis, bacteraemia
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
 CC region and the immediate 5' tip of the coding sequence are likely to be
 CC the critical sequences for predicting PspA cross-reactions and vaccine
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)

XX
 SQ Sequence 206 AA;
 Query Match 93.8%; Score 454; DB 2; Length 206;
 Best Local Similarity 92.9%; Pred. No. 7.3e-36;
 Matches 92; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 LEEINSDSDYAKGFRAPLQSKLDAKAKLLKLELSGKIELDIAEIAELEVQLKDAE 60
 Db 1 LKIDIDSDSDYAKGFRAPLQSKLDAKAKLLKLELSGKIELDIAEIAELEVQLKDAE 60
 Qy 61 GNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 99
 Db 61 GNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 2
 ABW02608
 ID ABW02608 standard; protein; 206 AA.

XX
 AC ABW02608;

DT 12-FEB-2004 (first entry)

XX Db15c pneumococcal surface protein A (PspA) central region.

DE Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
 KW immunological; gene therapy; immunostimulant.

XX Unidentified.

XX Key Location/Qualifiers
 FH Misc-difference 1. .206
 FT /note= "Xaa = Unknown amino acid"

XX US6592876-B1.

XX 15-JUL-2003.

XX 15-SEP-1995; 95US-00529055.

XX 20-APR-1993; 93US-00048896.

XX 06-JUN-1995; 95US-00465746.

XX (UABR-) UAB RES FOUND.

XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;

XX WPI; 2003-862841/80.

XX Immunological composition for obtaining expression products used for
 PT detecting the presence of Streptococcus pneumoniae or its strain,
 PT comprises at least two different full length isolated gene encoding
 PT pneumococcal surface protein A.

XX Example 6; SEQ ID NO 54; 121pp; English.

XX The present invention relates to an immunological composition comprising
 CC at least 2 different full length isolated genes encoding pneumococcal
 CC surface protein A (PspAs) from different groups based on restriction
 CC fragment polymorphism analysis. The invention is useful for obtaining
 CC expression products by recombinant techniques to detect, determine,
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its
 CC strain. The expression product is useful for preparing antigenic,
 CC immunological or vaccine compositions, for eliciting antibodies, an
 CC immunological response (other than or additional to antibodies) or a
 CC protective response (including antibody or other immunological response
 CC by administering compositions to a host). The invention is also useful as

CC vaccines and in gene therapy. The present sequence is Db15c pneumococcal
 CC surface protein A (PspA) central region. This sequence is used in the
 CC exemplification of the invention

XX Sequence 206 AA;

Query Match 93.8%; Score 454; DB 7; Length 206;
 Best Local Similarity 92.9%; Pred. No. 7.3e-36;
 Matches 92; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LEEINSDSDYAKGFRAPLQSKLDAKAKLLKLELSGKIELDIAEIAELEVQLKDAE 60
 Db 1 LKIDIDSDSDYAKGFRAPLQSKLDAKAKLLKLELSGKIELDIAEIAELEVQLKDAE 60

Qy 61 GNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 99

Db 61 GNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 3
 ABU08487
 ID ABU08487 standard; protein; 8991 AA.

XX
 AC ABU08487;

DT 24-JUN-2003 (first entry)

XX S. pneumoniae pneumococcal surface protein A (PspA) protein.

DE Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
 KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
 KW antibacterial.

XX Streptococcus pneumoniae.

XX Key Location/Qualifiers
 FH Misc-difference 1. .8991
 FT /note= "All Xaa residues within this sequence are
 FT unknown"

XX US6500613-B1.

XX 31-DEC-2002.

XX 16-SEP-1996; 96US-00714741.

XX 15-SEP-1995; 95US-00529055.

XX (UYAL-) UNIV ALABAMA.

XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;

XX Hollingshead S, Tart R, Brooks-Walter A;

XX WPI; 2003-361534/34.

XX Isolated PspC amino acid sequence used as polymerase chain reaction or
 PT hybridization probe, comprises pneumococcal surface protein having alpha-
 PT helical, proline rich and repeat regions.

XX Disclosure; Col 145-188; 186pp; English.

XX The present invention relates to the isolation of Streptococcus
 CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
 CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-
 CC like protein having alpha-helical, proline rich and repeat regions. The
 CC PspC and PspA proteins may be used in a vaccine to protect against
 CC pneumococcal infections. The polynucleotide sequences encoding PspC and
 CC PspA may be used for the expression of the proteins, and as PCR primers
 CC or hybridisation probes. The present sequence represents S. pneumoniae
 CC PspA protein

XX Sequence 8991 AA;

SQ

Db	1	LKEIDSDSDYLYKEGRAPLQSKLDTKAKLSKLEELSDKIDELDAETAKLEVQLKDAE	60
Qy	61	GNNVVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE	99
Db	61	GNNVVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE	99
RESULT 5			
ABW02596	ID	ABW02596 standard; protein; 181 AA.	
XX	AC	ABW02596;	
XX	XX		
DT	12-FEB-2004	(first entry)	
XX	XX		
DE	0922134C	pneumococcal surface protein A (PspA) central region.	
XX	XX		
KW	Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;		
KW	immunological; gene therapy; immunostimulant.		
XX	Unidentified.		
XX	OS		
PN	US6592876-B1.		
XX	XX		
PD	15-JUL-2003.		
XX	XX		
PF	15-SEP-1995;	95US-00529055.	
XX	XX		
PR	20-APR-1993;	93US-00048896.	
PR	06-JUN-1995;	95US-00465746.	
XX	XX		
PA	(UABR-) UAB RES FOUND.		
XX	XX		
PI	Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;		
XX	XX		
DR	WPI; 2003-862841/80.		
XX	XX		
PT	Immunological composition for obtaining expression products used for		
PT	detecting the presence of Streptococcus pneumoniae or its strain,		
PT	comprises at least two different full length isolated gene encoding		
PT	pneumococcal surface protein A.		
XX	XX		
PS	Example 6; SEQ ID NO 42; 121pp; English.		
XX	XX		
CC	The present invention relates to an immunological composition comprising		
CC	at least 2 different full length isolated genes encoding pneumococcal		
CC	surface protein A (PspAs) from different groups based on restriction		
CC	fragment polymorphism analysis. The invention is useful for obtaining		
CC	expression products by recombinant techniques to detect, determine,		
CC	isolate or diagnose the presence of Streptococcus pneumoniae or its		
CC	strain. The expression product is useful for preparing antigenic,		
CC	immunological or vaccine compositions, for eliciting antibodies, an		
CC	immunological response (other than or additional to antibodies) or a		
CC	protective response (including antibody or other immunological response		
CC	by administering compositions to a host). The invention is also useful as		
CC	vaccines and in gene therapy. The present sequence is 0922134C		
CC	pneumococcal surface protein A (PspA) central region. This sequence is		
CC	used in the exemplification of the invention		
XX	XX		
SQ	Sequence 181 AA;		
Query Match			
Query	Similarity	90.9%; Score 440; DB 7; Length 181;	
Best	Local Similarity	90.9%; Pred. No. 1.4e-34;	
Matches	90; Conservative	4; Mismatches 5; Indels 0; Gaps 0;	
Qy	1	LEINSDSDYLYKEGRAPLQSKLDAKAKLLKLEELSGKIDELDAETAKLEVQLKDAE	60
Db	1	LKEIDSDSDYLYKEGRAPLQSKLDTKAKLSKLEELSDKIDELDAETAKLEVQLKDAE	60
Qy	61	GNNVVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE	99
Db	61	GNNVVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE	99

```

RESULT 6
ABU08489
ID ABU08489 standard; protein; 865 AA.
XX AC ABU08489;
XX DT 24-JUN-2003 (first entry)
XX DE S. pneumoniae pneumococcal surface protein C (PspC) protein.
XX KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
XX KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
XX KW antibacterial.
XX OS Streptococcus pneumoniae.
XX FH Key Location/Qualifiers
XX FT Peptide 1..37
XX FT Protein /label= Signal_peptide
XX FT Protein 38..865
XX FT Protein /label= Mature_PspC_protein
XX US6500613-B1.
XX PN 31-DEC-2002.
XX PD 16-SEP-1996; 96US-00714741.
XX PF 15-SEP-1995; 95US-00529055.
XX PR (UYAL-) UNIV ALABAMA.
XX PA Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
XX PI Hollingshead S, Tart R, Brooks-Walter A;
XX DR WPI; 2003-361534/34.
XX DR N-PSDB; ABX95377.
XX PT Isolated PspC amino acid sequence used as polymerase chain reaction or
XX PT hybridization probe, comprises pneumococcal surface protein having alpha-
XX PT helical, proline rich and repeat regions.
XX PS Claim 3; Fig 21; 186pp; English.
XX CC The present invention relates to the isolation of Streptococcus
XX CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
XX CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-
XX CC like protein having alpha-helical, proline rich and repeat regions. The
XX CC PspC and PspA proteins may be used in a vaccine to protect against
XX CC pneumococcal infections. The polynucleotide sequences encoding PspC and
XX CC PspA may be used for the expression of the proteins, and as PCR primers
XX CC or hybridisation probes. The present sequence represents S. pneumoniae
XX CC PspC protein
XX SQ Sequence 865 AA;

Query Match 90.9%; Score 440; DB 6; Length 865;
Best Local Similarity 90.9%; Pred. No. 9.4e-34;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LEEINESDSEYAKGFRAPLQSKLDAAKALKLELSGKIELDIAEIAELEVQKDAE 60
DB 466 LKEIDESDSEYALKEGLRAPLQSKLDYTKAKLSKLELSKIDELDAEIAKLEVQKDAE 525

QY 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDPE 99
DB 526 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDPE 564

RESULT 7
AAW14593
ID AAW14593 standard; protein; 929 AA.
XX AC AAW14593;
XX DT 17-OCT-2003 (revised)
XX DT 27-OCT-1997 (first entry)
XX DE Streptococcus pneumoniae PspC.
XX KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
XX KW bacteraemia; pneumonia.
XX OS Streptococcus pneumoniae; strain EF6796.
XX FH Key Location/Qualifiers
XX FT Peptide 1..37
XX FT Protein /label= Sig_peptide
XX FT Protein 38..929
XX FT Domain /label= Mat_protein
XX FT Domain 38..637
XX FT Region /label= Alpha-helix
XX FT Region 41..242
XX FT /label= Repeat_1
XX FT /note= "alpha-helical repeat region"
XX FT 69..637
XX FT /label= Coiled-coil
XX FT /note= "breaks in 7-residue periodicity of the coiled-
XX FT coil occur at amino acids 136-141, 261-304 and 383-387"
XX FT 332..492
XX FT /label= Repeat_2
XX FT /note= "alpha-helical repeat region"
XX FT 627..689
XX FT /label= Proline-rich
XX FT 913..929
XX FT /label= C-terminal
XX WO9709994-A1.
XX PN 20-MAR-1997.
XX PD 16-SEP-1996; 96WO-US014819.
XX PF 15-SEP-1995; 95US-00529055.
XX PR (UABR-) UAB RES FOUND.
XX PA Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
XX PI Hollingshead S, Tart R, Brooks-Walter A;
XX DR WPI; 1997-202002/18.
XX DR N-PSDB; AAT61728.
XX PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used
XX PT in vaccines for protecting animals against S.pneumoniae infection.
XX PS Disclosure; Fig 13; 296pp; English.
XX CC This sequence comprises the pneumococcal protein surface C (pspC) of
XX CC Streptococcus pneumoniae strain EF6796. The sequence was deduced from the
XX CC pspC gene (AAT61728). Like PspA, PspC has an alpha-helical coiled-coil
XX CC region, proline-rich central region, repeat regions, with a choline
XX CC binding motif, and a C-terminal 17-amino acid tail. The 2 polypeptides
XX CC share 3 regions of high sequence identity. One is a protection-eliciting
XX CC region present within the alpha-helical domain. The others are the
XX CC proline-rich domain and a repeat domain shared with other choline-binding
XX CC proteins and thought to play a role in cell surface association. PspC and
XX CC PspA polypeptides, and their fragments, can be used in vaccines to
XX CC protect against S. pneumoniae infection and hence for the prevention of
XX CC diseases such as otitis media, meningitis, bacteraemia and pneumonia.
XX CC (Updated on 17-OCT-2003 to standardise OS field)
XX SQ Sequence 929 AA;

Query Match 90.9%; Score 440; DB 2; Length 929;

```

```
Best Local Similarity 90.9%; Pred. No. 1e-33; Mismatches 4; Indels 5; Gaps 0;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LBEINESDSEYAKGFRAPLQSKLDAAKAKLLKLELSGKIBELDAEIAEVLQKDAE 60
DB 530 LKDESDSEYDLKEGLRAPLQSKLDTKAKLKLSELSKIDELDAEIAKLEVLQKDAE 589

QY 61 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99
DB 590 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 628

RESULT 8
AAV43384
ID AAY43384 standard; protein; 929 AA.
XX
AC AAY43384;
XX
DT 27-JAN-2000 (first entry)
XX
DE S. pneumoniae PspC protein sequence.
XX
KW PspC gene; pneumococcal surface protein C; epitope; diagnosis;
KW epitopic region; immunogenic composition; vaccine composition; therapy;
KW meningitis; immunological response; otitis media; bacteraemia; pneumonia;
KW anti-PspA antibody.
XX
OS Streptococcus pneumoniae.
XX
PN WO9953940-A1.
XX
PD 28-OCT-1999.
XX
PF 23-APR-1999; 99WO-US008895.
XX
PR 23-APR-1998; 98US-0082728P.
XX
PA (UYAL-) UNIV ALABAMA.
XX
PI Briles DE, Hollingshead SK, Brooks-Walter A;
XX
DR WPI; 1999-620581/53.
XX
DR N-PSDB; AAZ31956.
XX
New epitope polypeptides of pneumococcal surface protein C, used to
PT develop products for immunological, immunogenic or vaccine compositions,
PT particularly against Streptococcus pneumoniae infections.
XX
PS Example; Fig 11; 109pp; English.
XX
This sequence is the Streptococcus pneumoniae pneumococcal surface
CC protein C. The invention relates to an isolated and/or purified
CC polypeptide comprising at least one epitope or epitopic region of
CC pneumococcal surface protein C (PspC). The polypeptides or vectors
CC containing sequence encoding them can be used as immunogenic,
CC immunological or vaccine compositions. The compositions can be used for
CC eliciting an immunological response against Streptococcus pneumoniae
CC (SP), which can cause otitis media, meningitis, bacteraemia and
CC pneumonia. They can be used for eliciting an anti-PspA antibody. The
CC nucleic acid molecules can also be used for detecting pspC, pspA or SP
XX
SQ Sequence 929 AA;

Query Match 90.9%; Score 440; DB 2; Length 929;
Best Local Similarity 90.9%; Pred. No. 1e-33;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LBEINESDSEYAKGFRAPLQSKLDAAKAKLLKLELSGKIBELDAEIAEVLQKDAE 60
DB 530 LKDESDSEYDLKEGLRAPLQSKLDTKAKLKLSELSKIDELDAEIAKLEVLQKDAE 589

QY 61 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99
DB 590 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 628
```

```
DB 590 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 628

RESULT 9
AAW14580
ID AAW14580 standard; protein; 188 AA.
XX
AC AAW14580;
XX
DT 17-OCT-2003 (revised)
DT 28-OCT-1997 (first entry)
XX
DE Streptococcus pneumoniae PspA central region.
XX
KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW bacteraemia; pneumonia.
XX
OS Streptococcus pneumoniae; strain Rct135.
XX
PN WO9709994-A1.
XX
PD 20-MAR-1997.
XX
PF 16-SEP-1996; 96WO-US014819.
XX
PR 15-SEP-1995; 95US-00529055.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Briles DE, Medaniel LS, Swiatlo E, Yother J, Crain MJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX
DR WPI; 1997-202002/18.
XX
Streptococcus pneumoniae surface protein PspC and truncated PspA - used
PT in vaccines for protecting animals against S.pneumoniae infection.
XX
PS Example 6; Fig 13; 296pp; English.
XX
This sequence shows the central portion, including the C-terminus of the
CC alpha-helix region and some of the proline-rich region, of pneumococcal
CC surface protein A (PspA) of Streptococcus pneumoniae strain Rct135.
CC Comparison of the N-terminal and central regions (AAW14533-57 and
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
CC be used to divide the strains into several families based on sequence
CC homologies. PspA polypeptides, or fragments of them, can be used in
CC vaccines to protect animals against S. pneumoniae infection and hence for
CC the prevention of diseases such as otitis media, meningitis, bacteraemia
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
CC region and the immediate 5' tip of the coding sequence are likely to be
CC the critical sequences for predicting PspA cross-reactions and vaccine
CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 188 AA;

Query Match 90.3%; Score 437; DB 2; Length 188;
Best Local Similarity 89.9%; Pred. No. 2.9e-34;
Matches 89; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LBEINESDSEYAKGFRAPLQSKLDAAKAKLLKLELSGKIBELDAEIAEVLQKDAE 60
DB 530 LKDESDSEYDLKEGLRAPLQSKLDTKAKLKLSELSKIDELDAEIAKLEVLQKDAE 60

QY 61 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99
DB 61 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPD 99

RESULT 10
ABW02613
ID ABW02613 standard; protein; 188 AA.
XX
AC ABW02613;
```



```

PR 15-SEP-1995; 95US-00529055.
XX (UYAL-) UNIV ALABAMA.
XX
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX
XX WPI; 2003-361534/34.
XX
XX Isolated PspC amino acid sequence used as polymerase chain reaction or
XX hybridization probe, comprises pneumococcal surface protein having alpha-
XX helical, proline rich and repeat regions.
XX
XX Disclosure; Fig 23; 186pp; English.
XX
XX The present invention relates to the isolation of Streptococcus
XX pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
XX sequence encoding it. PspC is a pneumococcal surface protein A (PspA) -
XX like protein having alpha-helical, proline rich and repeat regions. The
XX PspC and PspA proteins may be used in a vaccine to protect against
XX pneumococcal infections. The polynucleotide sequences encoding PspC and
XX PspA may be used for the expression of the proteins, and as PCR primers
XX or hybridisation probes. The present sequence represents a coiled coil
XX motif of the alpha-helix of S. pneumoniae PspC protein
XX
XX Sequence 588 AA;
XX
XX Query Match 88.4%; Score 428; DB 6; Length 588;
XX Best Local Similarity 90.7%; Pred No. 8.5e-33;
XX Matches 88; Conservative 4; Mismatches 5; Indels 0; Gaps 0
XX
Qy 1 LEEINESDSEDYAKGFRAPLQSKLDAAKAKLLKLEELSGRIEELDAIEAEVLQKDAE 60
Db 492 LKEIDESDSEDYLKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAIEAEVLQKDAE 551
Qy 61 GNNNVEAYFKEGLEKTTAEKKAELKABADLKKAUDE 97
Db 552 GNNNVEAYFKEGLEKTTAEKKAELKABADLKKAUDE 588
XX
RESULT 13
AAV43392
XX ID AAV43392 standard; protein; 589 AA.
XX
XX AAV43392;
XX
XX 27-JAN-2000 (first entry)
XX
XX PspC alpha-helix coiled-coil region.
XX
XX PspC gene; pneumococcal surface protein C; epitope; diagnosis;
XX epitopic region; immunogenic composition; vaccine composition; therapy;
XX meningitis; immunological response; otitis media; bacteraemia; pneumonia;
XX anti-PspA antibody.
XX
XX Streptococcus pneumoniae.
XX
XX WO9953940-A1.
XX
XX 28-OCT-1999.
XX
XX 23-APR-1999; 99WO-US008895.
XX
XX 23-APR-1998; 98US-0082728P.
XX
XX (UYAL-) UNIV ALABAMA.
XX
XX Briles DE, Hollingshead SK, Brooks-Walter A;
XX
XX WPI; 1999-620581/53.
XX
XX New epitope polypeptides of Pneumococcal surface protein C, used to
XX develop products for immunological, immunogenic or vaccine compositions.

```

CC alpha-helix region and some of the proline-rich region, of pneumococcal
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Rct123.
 CC Comparison of the N-terminal and central regions (AAW14533-57 and
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
 CC be used to divide the strains into several families based on sequence
 CC homologies. PspA polypeptides, or fragments of them, can be used in
 CC vaccines to protect animals against S. pneumoniae infection and hence for
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
 CC region and the immediate 5' tip of the coding sequence are likely to be
 CC the critical sequences for predicting PspA cross-reactions and vaccine
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)

XX
 SQ Sequence 204 AA;
 Query Match 88.0%; Score 426; DB 2; Length 204;
 Best Local Similarity 87.9%; Pred. No. 3.7e-33;
 Matches 87; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
 Qy 1 LEEINESDSYAKGFRAPLQSKLDAAKAKLLKLELSGKIELDIAEIAELEVQLKDAE 60
 Db 1 IKEXDESXSDYLKEGLRAPLQSKLDTTKAKLSKLELSDKIDELDAEIAKLEVLKDAE 60
 Qy 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
 Db 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 15

ABW02612
 ID ABW02612 standard; protein; 204 AA.

XX
 AC ABW02612;

DT 12-FEB-2004 (first entry)

XX Rct123c pneumococcal surface protein A (PspA) central region.

DE Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
 KW immunological; gene therapy; immunostimulant.

XX Unidentified.

XX Key Location/Qualifiers
 FH Misc-difference 1. .204
 FT /note= "Xaa = Unknown amino acid"

XX US6592876-B1.

XX 15-JUL-2003.

XX 15-SEP-1995; 95US-00529055.

XX 20-APR-1993; 93US-00048896.

PR 06-JUN-1995; 95US-00465746.

XX (UABR-) UAB RES FOUND.

XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;

XX WPI; 2003-862841/80.

XX Immunological composition for obtaining expression products used for
 PT detecting the presence of Streptococcus pneumoniae or its strain,
 PT comprises at least two different full length isolated gene encoding
 FT pneumococcal surface protein A.

XX Example 6; SEQ ID NO 58; 121bp; English.

XX The present invention relates to an immunological composition comprising
 CC at least 2 different full length isolated genes encoding pneumococcal
 CC surface protein A (PspAs) from different groups based on restriction
 CC fragment polymorphism analysis. The invention is useful for obtaining

CC expression products by recombinant techniques to detect, determine,
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its
 CC strain. The expression product is useful for preparing antigenic,
 CC immunological or vaccine compositions, for eliciting antibodies, an
 CC immunological response (other than or additional to antibodies) or a
 CC protective response (including antibody or other immunological response
 CC by administering compositions to a host). The invention is also useful as
 CC vaccines and in gene therapy. The present sequence is Rct123c
 CC pneumococcal surface protein A (PspA) central region. This sequence is
 CC used in the exemplification of the invention

XX SQ Sequence 204 AA;

Query Match 88.0%; Score 426; DB 7; Length 204;
 Best Local Similarity 87.9%; Pred. No. 3.7e-33;
 Matches 87; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
 Qy 1 LEEINESDSYAKGFRAPLQSKLDAAKAKLLKLELSGKIELDIAEIAELEVQLKDAE 60
 Db 1 IKEXDESXSDYLKEGLRAPLQSKLDTTKAKLSKLELSDKIDELDAEIAKLEVLKDAE 60
 Qy 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
 Db 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99

Search completed: June 21, 2005, 10:10:15

Job time : 74.8459 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:55:14 ; Search time 18.4867 Seconds
(without alignments)
399.760 Million cell updates/sec

Title: US-10-674-755-15

Perfect score: 484

Sequence: 1 LBEINESSEDYAKGFRAP.....KKAELKAEADLKAVDEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata1/iaa/5A COMB.pep:*
2: /cgn2_6/prodata1/iaa/5B COMB.pep:*
3: /cgn2_6/prodata1/iaa/6A COMB.pep:*
4: /cgn2_6/prodata1/iaa/6B COMB.pep:*
5: /cgn2_6/prodata1/iaa/PCTUS COMB.pep:*
6: /cgn2_6/prodata1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	484	100.0	99	4	US-09-147-875A-15
2	470	97.1	99	2	US-08-710-749-14
3	454	93.8	206	4	US-08-529-055-54
4	454	93.8	8991	4	US-08-714-741-32
5	451	93.2	99	4	US-09-147-875A-14
6	445	91.9	99	2	US-08-710-749-17
7	444.5	91.8	100	4	US-09-147-875A-10
8	440	90.9	99	2	US-08-710-749-13
9	440	90.9	170	4	US-08-529-055-60
10	440	90.9	181	4	US-08-529-055-42
11	440	90.9	864	4	US-08-714-741-40
12	437	90.3	99	4	US-09-147-875A-16
13	437	90.3	188	4	US-08-529-055-59
14	429	88.6	1231	4	US-08-714-741-41
15	428	88.4	141	4	US-09-286-981B-2
16	428	88.4	588	4	US-08-714-741-42
17	426	88.0	99	2	US-08-710-749-15
18	426	88.0	204	4	US-08-529-055-58
19	401	82.9	99	2	US-08-710-749-11
20	401	82.9	198	4	US-08-529-055-61
21	401	82.9	619	1	US-08-465-746-2
22	401	82.9	619	1	US-08-214-164-2
23	401	82.9	619	2	US-08-467-852A-3
24	401	82.9	619	2	US-08-246-636-2
25	401	82.9	619	2	US-08-247-491A-3
26	401	82.9	619	2	US-08-319-795-2
27	401	82.9	619	2	US-08-468-985-2

```

28 401 82.9 619 3 US-08-312-949-2 Sequence 2, Appli
29 401 82.9 648 1 US-08-072-070-2 Sequence 2, Appli
30 401 82.9 648 1 US-08-469-434-2 Sequence 2, Appli
31 401 82.9 648 1 US-08-214-222-2 Sequence 2, Appli
32 401 82.9 648 2 US-08-467-852A-2 Sequence 2, Appli
33 401 82.9 648 2 US-08-468-718-2 Sequence 2, Appli
34 401 82.9 648 2 US-08-247-491A-2 Sequence 2, Appli
35 401 82.9 648 3 US-08-446-201-3 Sequence 3, Appli
36 401 82.9 695 1 US-08-127-499A-23 Sequence 23, Appl
37 401 82.9 695 1 US-08-482-847-23 Sequence 23, Appl
38 393 81.2 99 2 US-08-710-749-10 Sequence 10, Appl
39 393 81.2 99 4 US-09-147-875A-11 Sequence 11, Appl
40 393 81.2 204 4 US-08-529-055-51 Sequence 51, Appl
41 389 80.4 288 3 US-08-312-949-4 Sequence 4, Appli
42 389 80.4 288 3 US-08-446-201-4 Sequence 4, Appli
43 387.5 80.1 100 4 US-09-147-875A-12 Sequence 12, Appl
44 378.5 78.2 289 1 US-08-072-070-4 Sequence 4, Appli
45 378.5 78.2 289 1 US-08-469-434-4 Sequence 4, Appli

```

ALIGNMENTS

```

RESULT 1
US-09-147-875A-15
; Sequence 15, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-15

```

```

Query Match      100.0%; Score 484; DB 4; Length 99;
Best Local Similarity 100.0%; Pred. No. 1.9e-40;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LBEINESSEDYAKGFRAPLOSKLDKAKKLLKLEELSGKIEELDAETAELEVLQKDAE 60
DB 1 LBEINESSEDYAKGFRAPLOSKLDKAKKLLKLEELSGKIEELDAETAELEVLQKDAE 60

QY 61 GNNVVEAYPKGLEKTTAEKKAELKAEADLKAVDEPE 99
DB 61 GNNVVEAYPKGLEKTTAEKKAELKAEADLKAVDEPE 99

```

```

RESULT 2
US-08-710-749-14
; Sequence 14, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/710,749
FILING DATE: 20-SEP-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454312-2074
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 99 amino acids
TYPE: amino acid
STRANDEDNESS: n/a
TOPOLOGY: linear
MOLECULE TYPE: amino acid
US-08-710-749-14

Query Match 97.1%; Score 470; DB 2; Length 99;
Best Local Similarity 97.0%; Pred. No. 4.5e-39;
Matches 96; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LEEINESDSYAKGPRAPLQSKLDKAKKALKLELSKIBELDAEIAELEVLKDAE 60
Db 1 LEEINESDSYAKGPRAPLQSKLDKAKKALKLELSKIBELDAEIAELEVLKDXVE 60
Qy 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 3

US-08-529-055-54
Sequence 54, Application US/08529055
Patent No. 6592876
GENERAL INFORMATION:
APPLICANT: Briles, David E.
APPLICANT: McDaniel, Larry S.
APPLICANT: Swiatlo, Edwin
APPLICANT: Yother, Janet
APPLICANT: Brooks-Walter, Alexis
TITLE OF INVENTION: Pneumococcal Genes, Portions
TITLE OF INVENTION: Thereof, Expression Products
TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
TITLE OF INVENTION: Portions and Products
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford, P.C.
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/529,055
FILING DATE: 15-SEP-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454312-2400
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 206 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-529-055-54

Query Match 93.8%; Score 454; DB 4; Length 206;
Best Local Similarity 92.9%; Pred. No. 4.1e-37;
Matches 92; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LEEINESDSYAKGPRAPLQSKLDKAKKALKLELSKIBELDAEIAELEVLKDAE 60
Db 1 LKIDIDESDYAKGPRAPLQSELPTTKAKLKLLELSKIBELDAEIAELEVLKDAE 60
Qy 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 4

US-08-714-741-32
Sequence 32, Application US/08714741
Patent No. 6500613
GENERAL INFORMATION:
APPLICANT: Briles, David E.
APPLICANT: McDaniel, Larry S.
APPLICANT: Swiatlo, Edwin
APPLICANT: Yother, Janet
APPLICANT: Crain, Marilyn J.
APPLICANT: Hollingshead, Susan
APPLICANT: Tart, Rebecca
APPLICANT: Brooks-Walter, Alexis
TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
TITLE OF INVENTION: PORTIONS AND PRODUCTS
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford, P.C.
STREET: 530 Fifth Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714,741
FILING DATE: 16-SEP-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Frommer Esq., William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454312-2460
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 8991 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: amino acid
US-08-714-741-32

```

Query Match          93.8%; Score 454; DB 4; Length 8991;
Best Local Similarity 92.9%; Pred.No.3.9e-35;
Matches 92; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY      1 LEEINESDSDYAKGFRAPLQSKLDAAKAKLLKEELSGKIIELEDAEIAELEVLQDKAE 60
       :.:|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      1KOIDEISDSDYAKGFRAPLQSELDTKKAKLLKEELSGKIIELEDAEIAELEVLQDKAE 5947

QY      61 GNNNVEAYPFKEGLEKTTTAEKKAELEKAEDLKKAVDPE 99
       |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      5948 GNNNVEAYPFKEGLEKTTTAEKKAELEKAEDLKKAVDPE 5986

RESULT 5
US-09-147-875A-14
; Sequence 14, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 99
; TYPE: PRP
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(99)
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid
US-09-147-875A-14
```

		Query Match	93.2%	Score 451	DB 4	Length 99;
		Best Local Similarity	92.9%	Pred. No.	3.3e-37;	
		Matches	9;	Mismatches	4; Indels	0; Gaps 0;
Qy	1	LEEINESDSBDYAKGPRAPLQSKLDAAKALLKEELSGKIEELDRAIFAEVLVKDAE	60			
		:::: :: :				
Dd	1	LKDIDESDSDYAKGGRAPLQSELDTKAALKLEELSGKIEELDAIIXEIVLVQLKDAE	60			
		:: :: ::				
Qy	61	GNNNVAYFKEGLEKTTAAEKKAELEKAEADLKKAVDPE	99			
		:: :: ::				
Dd	61	GNNNVAYFKEGLEKTTAAEKKAELEKAEADLKKAVDPE	99			
		:: :: ::				

RESULT 6
US-08-710-749-17
; Sequence 17, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710.749

FILING DATE: 20-SEP-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454312-2074
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 99 amino acids
TYPE: amino acid
STRANDEDNESS: n/a
TOPOLOGY: linear
MOLECULE TYPE: amino acid
US-08-710-749-17

	Query Match	91.9%	Score 445;	DB 2;	Length 99;
	Best Local Similarity	91.9%;	Pred. No. 1.3e-36;		
	Matches	91;	Conservative 4;	Mismatches 4;	Indels 0; Gaps 0;
Qy	1	LEEINSDSDYAKGFRAPLQSKLDAAKAKLLKLIELSGKIEELDAEIAELEVLQKDAE	60		
		: : : : : : : : : : : : : : : : : : :			
Dd	1	LKEIDESDSDYAKGFRAPLQSKLDAAKAKLSKLEELSCKIDELDAEIAKLEVLQKDAE	60		
		: : : : : : : : : : : : : : : : : : :			
Qy	61	GNNNVAYFKEGLEKTTAEKKAELEKAEDLKVADEPE	99		
		: : : : : : : : : : : : : : : : : : :			
Dd	61	GNNNVAYFKEGLEKTTAEKATSELEKAEDLKVADEPE	99		
		: : : : : : : : : : : : : : : : : : :			

```

RESULT 7
US-09-147-875A-10
; Sequence 10, Application US/09147875A
; Patent NO. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-10

```

	Query Match	91.8%; Score 444.5; DB 4;	Length 100;
	Best Local Similarity	93.0%;	Pred. No. 1.5e-36;
	Matches	93; Conservative 4;	Mismatches 2; Indels 1; Gaps 1;
Qy	1 LEEINSDSDYAKGFRAPLQSKLDAAKALKLLEELSGKIIEELDARIAELE-VQLKDA	59	 : : :
Dd	1 LKEIDESDSEDYAKGFRAPLQSKLDAAKALKLEELSDKIDELDAEIIAFLCEVQLKDA	60	 : : :
Qy	60 EGNNNVEAYFKEGLEKTATBKAALLEKAEADLKAVDEPE	99	 : : :
Dd	61 EGNNNVEAYFKEGLEKTATBKAALLEKAEADLKAVDEPE	100	 : : :

```

RESULT 8
US-08-710-749-13
; Sequence 13, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28

```

```

; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
US-08-529-055-60

Query Match 90.9%; Score 440; DB 4; Length 170;
Best Local Similarity 90.9%; Pred. No. 7.7e-36;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LEINSDSDYAKGFRAPLQSKLDPAKAKLKLKLELSGKIELDAETAELEVLKDAE 60
Db 1 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLELSDKIDELDAETAELEVLKDAE 60

QY 61 GNNNVAYPKFEGLEKTTAEKKALEKAEADLKKAVDEPE 99
Db 61 GNNNVAYPKFEGLEKTTAEKKALEKAEADLKKAVDEPE 99

RESULT 10
US-08-529-055-42
; Sequence 42, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 181 amino acids

```

```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-42

Query Match          90.9%; Score 440; DB 4; Length 181;
Best Local Similarity 90.9%; Pred. No. 8.3e-36;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LEEINESDSYAKGFRAPLQSKLDAAKAKLLKEELSGKIEELDAEIAELEVQLKDAE 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 LKEIDESDSYDLKEGRAPLQSKLDTKYKAKLKEELSDKIDELDAEIAELEVQLKDAE 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 GNNNVEAYFKEGLEKTTTAEKKAELKAEADLKKAVDPE 99
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 GNNNVEAYFKEGLEKTTTAEKKAELKAEADLKKAVDPE 99
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 11
US-08-714-741-40
; Sequence 40, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 864 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-714-741-40

Query Match          90.9%; Score 440; DB 4; Length 864;
Best Local Similarity 90.9%; Pred. No. 5.5e-35;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LEEINESDSYAKGFRAPLQSKLDAAKAKLLKEELSGKIEELDAEIAELEVQLKDAE 60

```

```
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-59

Query Match          90.3%; Score 437; DB 4; Length 188;
Best Local Similarity 89.9%; Pred. No. 1.7e-35;
Matches 89; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LEEINESDESYAKGFRAPLQSKLDKAKKALLKLEELSGKIEELDAEIAELEVLQKDAE 60
Db 1 LKEIDESDESYLKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLQKDAE 60

Qy 61 GNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 61 GNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDBPD 99

RESULT 14
US-08-714-741-41
; Sequence 41, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: EXPRESSION AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1231 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid

US-08-714-741-41
Query Match          88.6%; Score 429; DB 4; Length 1231;
Best Local Similarity 90.7%; Pred. No. 1e-33;
Matches 88; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 3 EINESDESYAKGFRAPLQSKLDKAKKALLKLEELSGKIEELDAEIAELEVLQKDAE 62
Db 494 EVQLSESESYAKGFRAPLQSKLDKAKKALLKLEELSDKIDELDAEIAKLEQLKDAE 553

Qy 63 NNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 554 NNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 590

RESULT 15
US-09-286-981B-2
; Sequence 2, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizenann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; CURRENT FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-286-981B-2

Query Match          88.4%; Score 428; DB 4; Length 141;
Best Local Similarity 90.7%; Pred. No. 9.2e-35;
Matches 88; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LEEINESDESYAKGFRAPLQSKLDKAKKALLKLEELSGKIEELDAEIAELEVLQKDAE 60
Db 45 LKEIDESDESYLKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLQKDAE 104

Qy 61 GNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDE 97
Db 105 GNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDE 141

Search completed: June 21, 2005, 10:25:21
Job time : 19.4867 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 10:12:15 ; Search time 63.2388 Seconds
(without alignments)
601.118 Million cell updates/sec

Title: US-10-674-755-15

Perfect score: 484

Sequence: 1 LEEINSESDYAKGFRAP.....KKAELKAEADLKKAVIDEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	484	100.0	99	15	US-10-674-755-15
2	454	93.8	206	15	US-10-299-636-69
3	451	93.2	99	15	US-10-674-755-14
4	444.5	91.8	100	15	US-10-674-755-10
5	440	90.9	170	15	US-10-299-636-75
6	440	90.9	181	15	US-10-299-636-57
7	440	90.9	643	15	US-10-299-636-95
8	440	90.9	670	9	US-09-748-875-63
9	440	90.9	670	10	US-09-298-523B-63
10	440	90.9	670	10	US-09-748-875-61
11	440	90.9	690	10	US-09-298-523B-61

12	440	90.9	691	9	US-09-748-875-1	Sequence 1, Appli
13	440	90.9	691	10	US-09-298-523B-1	Sequence 1, Appli
14	440	90.9	701	9	US-09-748-875-62	Sequence 62, Appl
15	440	90.9	701	10	US-09-298-523B-62	Sequence 62, Appl
16	440	90.9	707	9	US-09-748-875-2	Sequence 2, Appli
17	440	90.9	707	10	US-09-298-523B-2	Sequence 2, Appli
18	440	90.9	711	9	US-09-748-875-3	Sequence 3, Appli
19	440	90.9	711	10	US-09-298-523B-3	Sequence 3, Appli
20	440	90.9	739	17	US-10-732-923-3294	Sequence 3294, Ap
21	440	90.9	929	9	US-09-748-875-60	Sequence 60, Appl
22	440	90.9	929	10	US-09-298-523B-60	Sequence 60, Appl
23	440	90.9	929	15	US-10-299-636-94	Sequence 94, Appl
24	437	90.3	99	15	US-10-674-755-16	Sequence 16, Appl
25	437	90.3	188	15	US-10-299-636-74	Sequence 74, Appl
26	428	88.4	141	14	US-10-254-995-2	Sequence 2, Appli
27	428	88.4	589	9	US-09-748-875-14	Sequence 14, Appl
28	428	88.4	589	10	US-09-298-523B-14	Sequence 14, Appl
29	428	88.4	589	15	US-10-299-636-97	Sequence 97, Appl
30	426	88.0	204	15	US-10-299-636-73	Sequence 73, Appl
31	401	82.9	198	15	US-10-299-636-76	Sequence 76, Appl
32	401	82.9	354	15	US-10-299-636-105	Sequence 105, App
33	401	82.9	588	15	US-10-299-636-96	Sequence 96, Appl
34	401	82.9	619	10	US-09-882-774-1	Sequence 1, Appli
35	401	82.9	619	15	US-10-282-122A-73702	Sequence 73702, A
36	401	82.9	619	16	US-10-414-532-72	Sequence 72, Appl
37	393	81.2	99	15	US-10-674-755-11	Sequence 11, Appl
38	393	81.2	204	15	US-10-299-636-66	Sequence 66, Appl
39	387.5	80.1	100	15	US-10-674-755-12	Sequence 12, Appl
40	377	77.9	193	15	US-10-299-636-64	Sequence 64, Appl
41	374	77.3	195	15	US-10-299-636-86	Sequence 86, Appl
42	373	77.1	99	15	US-10-674-755-17	Sequence 17, Appl
43	370	76.4	99	15	US-10-674-755-13	Sequence 13, Appl
44	353.5	73.0	336	15	US-10-299-636-103	Sequence 103, App
45	305.5	63.1	183	15	US-10-299-636-65	Sequence 65, Appl

ALIGNMENTS

RESULT 1

US-10-674-755-15
; Sequence 15, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-15

Query Match 100.0%; Score 484; DB 15; Length 99;
Best Local Similarity 100.0%; Pred. No. 1.5e-33;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	LEEINSESDYAKGFRAPLQSKLDKAKKLLKLELSGKIELDAIEAEVQLKDAE	60
Db	1	LEEINSESDYAKGFRAPLQSKLDKAKKLLKLELSGKIELDAIEAEVQLKDAE	60

Qy	61	GNNVVEAYFKGLEKTTAKKAELEKAEADLKKAVIDEPE	99
Db	61	GNNVVEAYFKGLEKTTAKKAELEKAEADLKKAVIDEPE	99

RESULT 2

000 CARPENTARIACONVULSANTREMBLANT 090

RESULT 9
US-09-298-523B-63
; Sequence 63, Application US/09298523B
; Publication No. US20030059438A1
; GENERAL INFORMATION:
; APPLICANT: BRILLES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/298,523B
; CURRENT FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 670
; TYPE: PRF
; ORGANISM: Streptococcus pneumoniae

US-09-298-523B-63

Query Match 90.9%; Score 440; DB 10; Length 670;
Best Local Similarity 90.9%; Pred. No. 6.7e-29;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
Qy 1 LEEINESDSEDYAKGFRAPLQSKLDKAKKLLKLELSGKIELDIAEIALEVLQKDAE 60
Db 498 LKEIDESDSEDYKGLRAPLQSKLDTKAKLKLKLELSGKIELDIAEIALEVLQKDAE 557
Qy 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 558 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 596

RESULT 10

US-09-748-875-61
; Sequence 61, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; PRIOR FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 61
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-61

Query Match 90.9%; Score 440; DB 9; Length 690;
Best Local Similarity 90.9%; Pred. No. 6.9e-29;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
Qy 1 LEEINESDSEDYAKGFRAPLQSKLDKAKKLLKLELSGKIELDIAEIALEVLQKDAE 60
Db 529 LKEIDESDSEDYKGLRAPLQSKLDTKAKLKLKLELSGKIELDIAEIALEVLQKDAE 588
Qy 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 589 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 627

RESULT 11

US-09-298-523B-61
; Sequence 61, Application US/09298523B
; Publication No. US20030059438A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/298,523B
; CURRENT FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 61
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-61

Query Match 90.9%; Score 440; DB 10; Length 690;
Best Local Similarity 90.9%; Pred. No. 6.9e-29;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
Qy 1 LEEINESDSEDYAKGFRAPLQSKLDKAKKLLKLELSGKIELDIAEIALEVLQKDAE 60

Db 529 LKEIDESDSEDYKGLRAPLQSKLDTKAKLKLKLELSGKIELDIAEIALEVLQKDAE 588
Qy 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 589 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 627

RESULT 12

US-09-748-875-1
; Sequence 1, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; PRIOR FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-1

Query Match 90.9%; Score 440; DB 9; Length 691;
Best Local Similarity 90.9%; Pred. No. 6.9e-29;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
Qy 1 LEEINESDSEDYAKGFRAPLQSKLDKAKKLLKLELSGKIELDIAEIALEVLQKDAE 60
Db 530 LKEIDESDSEDYKGLRAPLQSKLDTKAKLKLKLELSGKIELDIAEIALEVLQKDAE 589
Qy 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 590 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 628

RESULT 13

US-09-298-523B-1
; Sequence 1, Application US/09298523B
; Publication No. US20030059438A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/298,523B
; CURRENT FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-1

Query Match 90.9%; Score 440; DB 10; Length 691;
Best Local Similarity 90.9%; Pred. No. 6.9e-29;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LEEINESDSEDYAKGFRAPLQSKLDKAKKLLKLELSGKIELDIAEIALEVLQKDAE 60
Db 530 LKEIDESDSEDYKGLRAPLQSKLDTKAKLKLKLELSGKIELDIAEIALEVLQKDAE 589
Qy 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 590 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 628

Search completed: June 21, 2005, 11:18:35
Job time : 63.2388 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:53:39 ; Search time 9.9 Seconds
(without alignments)
962.168 Million cell updates/sec

Title: US-10-674-755-15
Perfect score: 484
Sequence: 1 LEEINESDESDYAKGFRAP.....KKAELEKADLKKAVDEPE 99
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	401	82.9	619	2 A97887	surface protein ps
2	401	82.9	619	2 A41971	surface protein ps
3	121	25.0	744	2 P95013	pneumococcal surfa
4	120.5	24.9	161	2 S48396	tropomyosin TPM2 -
5	114.5	23.7	852	2 D72230	conserved hypothet
6	113.5	23.5	1006	2 C70445	ATPase subunit of
7	110	22.7	233	2 S70531	bbk2.11 protein pr
8	108.5	22.4	281	2 F75216	hypothetical prote
9	102.5	21.2	650	2 A11333	ABC transporter (A
10	102	21.1	1319	2 A28313	glued protein - fr
11	101	20.9	369	2 AG1648	hypothetical prote
12	100.5	20.8	229	2 S70532	outer surface prot
13	100.5	20.8	2401	2 T28676	rhostry protein -
14	100	20.7	880	2 F75103	conserved hypothet
15	98.5	20.4	399	2 E71169	hypothetical prote
16	98.5	20.4	701	2 H98120	choline binding pr
17	98.5	20.4	785	2 T01025	hypothetical prote
18	98	20.2	1078	2 T18352	protein P120 - Msc
19	97.5	20.1	1110	2 I51116	NF-180 - sea lampr
20	97	20.0	1790	2 S67593	transport protein
21	96	19.8	395	2 AC1754	capsid protein (ba
22	96	19.8	650	2 AH1704	ABC transporter (A
23	95.5	19.7	764	2 T05409	hypothetical prote
24	94	19.4	635	2 A10625	ABC transporter AT
25	94	19.4	2116	2 A26655	myosin heavy chain
26	93.5	19.3	407	1 EDBRQ3	immediate-early pr
27	93.5	19.3	935	2 T51930	kinesin [imported]
28	93	19.2	2139	2 T18296	myosin heavy chain
29	92.5	19.1	488	2 F97039	hypothetical prote

30	92.5	19.1	886	2 H69378	conserved hypothet
31	92	19.0	1169	2 A64505	Flis homolog - Met
32	92	19.0	1173	2 T25539	hypothetical prote
33	92	19.0	1177	2 B75150	chromosome segrega
34	92	19.0	1269	2 F84730	probable myosin he
35	92	19.0	1827	2 T16270	hypothetical prote
36	91.5	18.9	403	2 T16171	hypothetical prote
37	91.5	18.9	522	2 C98608	hypothetical prote
38	91	18.8	646	2 AH1587	bacteriophage prot
39	91	18.8	660	2 AC1652	hypothetical prote
40	91	18.8	1392	2 A43336	microtubule-vesicl
41	91	18.8	1427	2 S22695	resin - human
42	90.5	18.7	419	2 G75062	probable flagella-
43	90	18.6	433	2 A89951	trigger factor lim
44	90	18.6	1133	2 T22976	hypothetical prote
45	90	18.6	1138	2 T24635	hypothetical prote

ALIGNMENTS

RESULT 1

A97887
surface protein pspA precursor [imported] - Streptococcus pneumoniae (strain R6)
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C;Accession: A97887
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: A97887
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-619 <KUR>
A;Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:AE007317; PIDN:AAK98925.1; PID:gi:
C;Genetics:
A;Gene: pspA

Query Match	82.9%;	Score	401;	DB	2;	Length	619;
Best Local Similarity	84.8%;	Pred. No.	7.9e-21;				
Matches	84;	Conservative	6;	Mismatches	9;	Indels	0;
Gaps	0;						
QY	1	LEEINESDESDYAKGFRAPLOS	KLDAKKAKLLKLELSGKIEELDAEIAELSVQLKDAE	60			
Db	223	LKEIDSESDYAKGFRAPLOS	KLDAKKAKLSKLESDKIDELDAEIAELSDQLKAAE	282			
QY	61	GNNVVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE	99				
Db	283	ENNVVEDYFKEGLEKTTAAKAELEKAEADLKKAVNEPE	321				

RESULT 2

A41971
surface protein pspA precursor - Streptococcus pneumoniae
N;Alternate names: pneumococcal surface protein A
C;Species: Streptococcus pneumoniae
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A41971; A60282; A33134
R;Yother, J.; Briles, D.E.
J. Bacteriol. 174, 601-609, 1992
A;Title: Structural properties and evolutionary relationships of PspA, a surface protein
A;Reference number: A41971; MUID:92105030; PMID:1729249
A;Accession: A41971
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-619 <YOT>
A;Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:M74122; NID:q153840; PIDN:AAA2701
A;Note: sequence extracted from NCBI backbone (NCBIN:75635, NCBIIP:75636)
R;Talkington, D.F.; Crammins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.


```
QY 62 NNNVEAYFKEGLEKTTAAKAELEKAEAD 90
Db 566 E-----TEAKRETLKEOREMDQLKSD 587

RESULT 6
C70445
ATPase subunit of ATP-dependent proteinase (EC 3.4.-.-) - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C:Accession: C70445
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: C70445
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1006 <AQF>
A:Cross-references: UNIPROT:O67588; GB:AE000750; NID:g2983999; PIDN:AA07550.1; PID:g298
A:Experimental source: strain VF5
C:Genetics:
A:Gene: clpB
C:Superfamily: endopeptidase Clp ATP-binding chain
C:Keywords: hydrolase

Query Match 23.5%; Score 113.5; DB 2; Length 1006;
Best Local Similarity 36.3%; Pred. No. 1.2;
Matches 37; Conservative 22; Mismatches 24; Indels 19; Gaps 5;

QY 2 EINESDSE-DYAKGFRAPLOSKLDKAKLLK-LEELSGKTEELDAEIAEVLKDA 59
Db 557 EQIIIEANLKG DYKE-----AQLKIEAKLEKEQELLGKGVGEAKIAELKKKIEB- 608

QY 60 EGNVNEAYFKEGLEKTTAAKAELE-----KAEADLKKAUDE 97
Db 609 -----LDEKIEAAERGDYKEAELEKIEAKLEKLEKLEQE 645

RESULT 7
C70531
bbk2.11 protein precursor - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S70531
R:Akins, D.R.; Porcella, S.F.; Popova, T.G.; Shevchenko, D.; Baker, S.I.; Li, M.; Norgard
Mol. Microbiol. 18, 507-520, 1995
A:Title: Evidence for in vivo but not in vitro expression of a Borrelia burgdorferi outer
A:Reference number: S70531; MUID:96342380; PMID:8748034
A:Accession: S70531
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-233 <AKI>
A:Cross-references: UNIPROT:Q44739; EMBL:U30617; NID:g3309515; PIDN:AA046421.1; PID:g119
C:Superfamily: outer surface protein P ospF
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-233/Product: bbk2.11 protein #status predicted <MAT>

Query Match 22.7%; Score 110; DB 2; Length 233;
Best Local Similarity 31.7%; Pred. No. 0.5;
Matches 38; Conservative 20; Mismatches 36; Indels 26; Gaps 5;

QY 3 EINESDEDYAK-----EGFRAPLO-----SKLDKAKLL--KLEELSGKTEELD 46
Db 33 EQNLESSEQNKKTBQIBKQVEGFULEITKDLKSLDEKDTKEIBKQIQELKNKIEKLD 92

QY 47 AEIAELEV-----QLKDAEGNNVNEAYFKEGLEKTTAAKAELEKAEADLKKAUDE 97
Db 93 SKTSTIETSYEYBEKINKIEKLKGKLEDKFKE-LEESLAKKKGKKALQAKQKFE 151

us-10-674-755-15.rpr Page 3
```

R;Akins, D.R.; Porcella, S.F.; Popova, T.G.; Shevchenko, D.; Baker, S.I.; Li, M.; Norgard Mol. Microbiol. 18, 507-520, 1995
A;Title: Evidence for in vivo but not in vitro expression of a Borrelia burgdorferi outer A;Reference number: S70531; MUID:96342380; PMID:8748034
A;Accession: S70532
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-229 <AKI>
A;Cross-references: UNIPROT:Q44735; EMBL:U19754; NID:g3318660; PIDN:AAC26147.1; PID:g896 C;Genetics:
A;Gene: ospF
C;Superfamily: outer surface protein F ospF
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-229/Product: outer surface protein F #status predicted <MAT>

Query Match 20.8%; Score 100.5; DB 2; Length 229;
Best Local Similarity 33.9%; Pred. No.2.3;
Matches 38; Conservative 20; Mismatches 35; Indels 19; Gaps 6;

Qy 3 EINESDSDYAK---EGFRAPLQ-----SKLDAAKAKLLKLLELSGKIIEELDAE-----I 49
Db :
40 EQNVKTEQEIKQVEGFLEILETKDLANKLDTKEIE-KRIQLKEKIEKLEAKTSLKTY 98

Qy 50 AELEVQLKD-----AEGNNNVAYFKEGLEKTATTAEKKAEELEKADLKKAVIDE 97
Db :
99 SEYBEKLQIKEKLKGKADLEDKLK-GLEDSLKKKKERKKALEDAKKKFEE 149

RESULT 13
T28676
rhostry protein - Plasmodium yoelii (fragment)
C;Species: Plasmodium yoelii
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T28676; A45521
R;Sinha, K.A.; Keen, J.K.; Ogun, S.A.; Holder, A.A.
Mol. Biochem. Parasitol. 76, 329-332, 1996
A;Title: Comparison of two members of a multigene family coding for high-molecular mass A;Reference number: Z20507; MUID:97077455; PMID:8920022
A;Accession: T28676
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2401 <SN>
A;Cross-references: UNIPROT:Q26216; EMBL:U36927; NID:g1041784; PID:g1041785; PIDN:AAAB412 R;Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.
Mol. Biochem. Parasitol. 42, 241-246, 1990
A;Title: Identification of the gene for a Plasmodium yoelii rhostry protein. Multiple co A;Reference number: A45521; MUID:91101660; PMID:2270106
A;Accession: A45521
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 2260-2401 <KEE>
A;Cross-references: GB:M34281

Query Match 20.8%; Score 100.5; DB 2; Length 2401;
Best Local Similarity 34.0%; Pred. No.23;
Matches 33; Conservative 14; Mismatches 45; Indels 5; Gaps 2;

Qy 6 ESDESDYAKEGFRAPLQSKLDAKAKLLKLLELSGKIEB---LDAEIAELEVLKDAEGN 62
Db :
165 DKSTYEDEKGEFSSELAKNWEKKLIITTELKKNEETVQLDIKIRELIKQIKDIEE 224

Qy 63 NNVEYAFKEGLEKTAAE--KKAEELEKAEADLKKAVIDE 97
Db :
225 QKVNDLKLNLKNIKEITEKIEYIKKAVDLKKEIEK 261

RESULT 14
F75103
conserved hypothetical protein PAB0812 - Pyrococcus abyssi (strain Orsay)
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: F75103
R;anonymous, Genoscope


```

submitted to the EMBL Data Library, July 1999
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A;Reference number: A75001
A;Accession: F75103
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-880 <KAW>
A;Cross-references: UNIPROT:Q9U2C8; GB:AJ248286; GB:AL096836; NID:G5458366; PIDN:CAB5013
A;Experimental source: strain Orsay
C;Genetics:
C;Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match      20.7%; Score 100; DB 2; Length 880;
Best Local Similarity 27.9%; Pred. No. 9.2;
Matches 34; Conservative 18; Mismatches 38; Indels 32; Gaps 3;

QY      1 L E I N S D S D Y A K E G F R A P L Q S K L D A K K A K L K L E E L S K I E E L D A E I A E L V Q L K D A E 60
Db      598 L E E F H K Y V E A K S E S E L R E L K N K L E K E T E L D Q A F E M L A D V E N --- E T E E K A E K L D L E 654

QY      61 G N N N V E A Y ----- F K E G L E K T T A ----- E K K A E L E K A E A D L 91
Db      655 S K F E E Y E B K R E R L V K L E R E V S S L T A R L E E L K S V E Q I K A T L R K L K E K E R E K A K L E I 714

QY      92 K K 93
Db      715 K K 716

RESULT 15
E71169
hypothetical protein PH0553 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 09-Jul-2004
C;Accession: E71169
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamanoto, S.; Sekin
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: E71169
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-399 <KAW>
A;Cross-references: UNIPROT:O58288; GB:AP000002; NID:G3236129; PIDN:BAA29642.1; PID:G328
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH0553

Query Match      20.4%; Score 98.5; DB 2; Length 399;
Best Local Similarity 27.8%; Pred. No. 5.4;
Matches 35; Conservative 20; Mismatches 40; Indels 31; Gaps 3;

QY      4 I N E S D S D Y A K E G F R A P ----- L Q S K L D A K K A K L K L E E L S K I E E L D A E I A E L V Q L - 56
Db      27 I N E L R E K L M A K D T L T P K Q L D E I I K R V L D A Y G S Q A A K Y E Q I S K R V D E L G K K L G E L S S Q L S 86

QY      57 ----- K D A E G N N N V E A Y F K E G L E K T T A E K K A E L E K A E A D L K K 93
Db      87 R L V E A L E E K F A V H E K A E S I A E K A A E V T E K V E R - I E E L L E E K P K E K S E L A K K V E I H K 145

QY      94 A V D E P E 99
Db      146 K V E E L E 151

```

Search completed: June 21, 2005, 10:12:01
Job time : 10.9 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:54:24 ; Search time 61.3194 Seconds
(without alignments)
826.751 Million cell updates/sec

Title: US-10-674-755-15
Perfect score: 484
Sequence: 1 LBEINESDEYAKGFRAP.....KKALEKAEADLUKKAVIDEPE 99

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_03: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	484	100.0	395	2 Q9LAY2	Q9lay2 streptococc
2	484	100.0	408	2 Q9LAY0	Q9lay0 streptococc
3	465	96.1	224	2 Q8GNS8	Q8gns8 streptococc
4	462	95.5	249	2 Q9L575	Q9l575 streptococc
5	459	94.8	426	2 Q9LAY5	Q9lay5 streptococc
6	452	93.4	99	2 Q8KQK4	Q8kqk4 streptococc
7	440	90.9	739	2 Q9RQT4	Q9rqt4 streptococc
8	440	90.9	820	2 Q9RQT1	Q9rqt1 streptococc
9	440	90.9	929	2 Q9KK19	Q9kk19 streptococc
10	440	90.9	929	2 Q9ZAY5	Q9zay5 streptococc
11	427	88.2	437	2 Q9LAY4	Q9lay4 streptococc
12	401	82.9	619	2 Q54972	Q54972 streptococc
13	401	82.9	619	2 Q8DR10	Q8dr10 streptococc
14	400.5	82.7	869	2 Q9KK27	Q9kk27 streptococc
15	393	81.2	417	2 Q9LAY3	Q9lay3 streptococc
16	379	78.3	415	2 Q9LAY1	Q9lay1 streptococc
17	312.5	64.6	225	2 Q9L591	Q9l591 streptococc
18	311.5	64.4	394	2 Q9LAY6	Q9lay6 streptococc
19	311.5	64.4	395	2 Q9LAZ1	Q9laz1 streptococc
20	309.5	63.9	246	2 Q9L578	Q9l578 streptococc
21	306.5	63.3	222	2 Q9L577	Q9l577 streptococc
22	306.5	63.3	262	2 Q9L576	Q9l576 streptococc
23	306.5	63.3	415	2 Q9LAY7	Q9lay7 streptococc
24	303.5	62.7	393	2 Q9LAZ3	Q9laz3 streptococc
25	301.5	62.3	194	2 Q9L5B5	Q9l5b5 streptococc
26	301.5	62.3	218	2 Q6UBB2	Q6ub2 streptococc
27	301.5	62.3	233	2 Q9L568	Q9l568 streptococc
28	301.5	62.3	236	2 Q9L569	Q9l569 streptococc
29	301.5	62.3	243	2 Q9L564	Q9l564 streptococc
30	301.5	62.3	243	2 Q9L567	Q9l567 streptococc
31	301.5	62.3	244	2 Q9L565	Q9l565 streptococc

RESULT 1

Q9LAY2 PRELIMINARY; PRT; 395 AA.
AC Q9LAY2;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF6796;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K.; Becker R.; Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071813; AAF2709.1; -;
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin_abc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 395 395
SQ SEQUENCE 395 AA; 42963 MW; 5886EF956BCBCC1E CRC64;

Query Match 100.0%; Score 484; DB 2; Length 395;
Best Local Similarity 100.0%; Pred. No. 2.5e-23;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LBEINESDEYAKGFRAPLOSKLDKAKKLLKLEELSGKTEELDAETAEVLKDAE 60
Db 225 LBEINESDEYAKGFRAPLOSKLDKAKKLLKLEELSGKTEELDAETAEVLKDAE 284
Qy 61 GNNVVEAYFKGLEKTTAEKKALEKAEADLUKKAVIDEPE 99
Db 285 GNNVVEAYFKGLEKTTAEKKALEKAEADLUKKAVIDEPE 323

RESULT 2

Q9LAY0 PRELIMINARY; PRT; 408 AA.
AC Q9LAY0;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

```

OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9163;
RX MEDLINE=20448953; PubMed=1092499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071815; AAF27711.1; -.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin_alc_DH_like.
FT NON_TER 408
FT NON_TER 408
SQ SEQUENCE 408 AA; 44254 MW; 4F64D874217297BF CRC64;

Query Match 100.0%; Score 484; DB 2; Length 408;
Best Local Similarity 100.0%; Pred. No. 2.5e-23;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LEEINESDSEDYAKGFRAPLQSLDKAKKLLKLEELSGKIEELDAEIAEVLKDAE 60
Db 228 LEEINESDSEDYAKGFRAPLQSLDKAKKLLKLEELSGKIEELDAEIAEVLKDAE 287

Qy 61 GNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 288 GNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 326

RESULT 3
Q8GNS8
ID Q8GNS8 PRELIMINARY; PRT; 224 AA.
AC Q8GNS8
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2004 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=PspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PH124;
RX MEDLINE=22241996; PubMed=12354862;
RA Dicuonzo G., Gherardi G., Gettz R.E., D'Ambrosio F., Goglio A.,
RA Lorino G., Recchia S., Pantosti A., Beall B.;
RT "Genotypes of invasive pneumococcal isolates recently recovered from
Italian patients.";
RL J. Clin. Microbiol. 40:3660-3665(2002).
DR EMBL; AF490267; AAN37735.1; -.
DR HSP; P00192; IAPC.
DR InterPro; IPR009082; His_kin_homodim.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 224
FT NON_TER 224
SQ SEQUENCE 224 AA; 23418 MW; 48674E27AFB66A95 CRC64;

Query Match 96.1%; Score 465; DB 2; Length 224;
Best Local Similarity 94.9%; Pred. No. 2.4e-22;
Matches 94; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LEEINESDSEDYAKGFRAPLQSLDKAKKLLKLEELSGKIEELDAEIAEVLKDAE 60
Db 17 LKIDINESDSEDYVKEGFRAPLQSLDTKKAALLKLEELSGKIEELDAEIAEVLKDAE 76

Qy 61 GNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 77 GNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 115

RESULT 4

```

```

Q9L575
ID Q9L575 PRELIMINARY; PRT; 249 AA.
AC Q9L575
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=PspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=195;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Packham R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=195;
RX Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255552; AAF68105.1; -.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 249
FT NON_TER 249
SQ SEQUENCE 249 AA; 26986 MW; 7916D5014E387BD8 CRC64;

Query Match 95.5%; Score 462; DB 2; Length 249;
Best Local Similarity 94.9%; Pred. No. 4e-22;
Matches 94; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LEEINESDSEDYAKGFRAPLQSLDKAKKLLKLEELSGKIEELDAEIAEVLKDAE 60
Db 74 LKIDINESDSEDYIKGFRAPLQSLDTKKAALLKLEELSGKIEELDAEIAEVLKDAE 133

Qy 61 GNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 134 GNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 172

RESULT 5
Q9L575
ID Q9L575 PRELIMINARY; PRT; 426 AA.
AC Q9L575
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=PspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBL5;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071810; AAF27706.1; -.
DR HSP; P00192; IM6T.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR00533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 426
FT NON_TER 426
SQ SEQUENCE 426 AA; 46534 MW; 81AA11348CBE6634 CRC64;

```

Gap 0;

```

Qy 1 LEEINESDSEYAKGFRAPLQSKLDKAKKLLKLEELSGKIBELDAEIAELEVLKDAE 60
Db 530 LKEIDESDSEYKKEGLRAPLQSKLDTKAKLSKLEELSDKIDELDAEIAELEVLKDAE 589

Qy 61 GNNNVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 99
Db 590 GNNNVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 628

RESULT 9
Q9KK19 PRELIMINARY; PRT; 929 AA.
AC Q9KK19;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein PspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=stf10;
RA MEDLINE=21888621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RX Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae.";
RL Gene 284:63-71(2002).
DR EMBL; AF154037; AAF73809.1; -.
DR HSSP; P06653; 1H8G.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk signal; 1.
SQ SEQUENCE 929 AA; 105003 MW; 2DC8293302FAFA64 CRC64;

Query Match 90.9%; Score 440; DB 2; Length 929;
Best Local Similarity 90.9%; Pred. No. 3.4e-20;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LEEINESDSEYAKGFRAPLQSKLDKAKKLLKLEELSGKIBELDAEIAELEVLKDAE 60
Db 530 LKEIDESDSEYKKEGLRAPLQSKLDTKAKLSKLEELSDKIDELDAEIAELEVLKDAE 589

Qy 61 GNNNVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 99
Db 590 GNNNVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 628

RESULT 10
Q9ZAY5 PRELIMINARY; PRT; 929 AA.
AC Q9ZAY5;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein C.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF6796;
RX MEDLINE=20038319; PubMed=10569772;

```

```

RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, PspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; U72655; AAD00184.1; -.
DR HSSP; P06653; 1HXC.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk signal; 1.
SQ SEQUENCE 929 AA; 104991 MW; 2DC8293302FFB081 CRC64;

Query Match 90.9%; Score 440; DB 2; Length 929;
Best Local Similarity 90.9%; Pred. No. 3.4e-20;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LEEINESDSEYAKGFRAPLQSKLDKAKKLLKLEELSGKIBELDAEIAELEVLKDAE 60
Db 530 LKEIDESDSEYKKEGLRAPLQSKLDTKAKLSKLEELSDKIDELDAEIAELEVLKDAE 589

Qy 61 GNNNVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 99
Db 590 GNNNVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 628

RESULT 11
Q9LAY4 PRELIMINARY; PRT; 437 AA.
AC Q9LAY4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E134;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071811; AAF27707.1; -.
DR InterPro; IPR002479; CW binding.
DR Pfam; PF01473; CW binding_1; 1.
DR NON_TER 437
SQ SEQUENCE 437 AA; 48071 MW; F6EFD2CD13E08CD8 CRC64;

Query Match 88.2%; Score 427; DB 2; Length 437;
Best Local Similarity 88.9%; Pred. No. 1.1e-19;
Matches 88; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 LEEINESDSEYAKGFRAPLQSKLDKAKKLLKLEELSGKIBELDAEIAELEVLKDAE 60
Db 235 LKEIDESDSEYKKEGLRAPLQSKLDTKAKLSKLEELSDKIDELDAEIAEHVVLKDAE 294

Qy 61 GNNNVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 99
Db 295 GNNNVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 333

RESULT 12

```

```

Q54972 PRELIMINARY; PRT; 619 AA.
AC Q54972;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pneumococcal surface protein A precursor.
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]_TaxID=1313;
RP SEQUENCE FROM N.A.
RX MEDLINE=92105030; PubMed=1729249;
RA Yother J., Briles D.E.;
RT "Structural properties and evolutionary relationships of PspA, a
RT surface protein of Streptococcus pneumoniae, as revealed by sequence
RT analysis."
RL J. Bacteriol. 174:601-609(1992).
[2]
RN SEQUENCE FROM N.A.
RA Yother J., Briles D.E.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M74122; AAA27018.1; -.
DR PIR; A41971; A41971.
DR PIR; A97887; A97887.
DR HSSP; P06653; 1HCX.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam; PF01473; CW_binding_1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_3.
KW Signal.
FT SIGNAL.
FT CHAIN 1 31 Potential.
FT CHAIN 32 619 pneumococcal surface protein A.
SQ SEQUENCE 619 AA; 68605 MW; 5AA8BDB40C2841CA CRC64;
Query Match 82.9%; Score 401; DB 2; Length 619;
Best Local Similarity 84.8%; Pred. No. 7.1e-18;
Matches 84; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 LEEINSESDYAKGFRAPLOSKLDKAKKLLKLELSGKIEELDAETAELEVLKDAE 60
DB 223 LKEIDSESDYAKGFRAPLOSKLDKAKKLLKLELSGKIEELDAETAELEVLKDAE 282

QY 61 GNNVYAYFKEGKLETTAEKAELEKAEADLKAVDEPE 99
DB 283 ENNVEDYFKEGLEKTTAAKAELEKAEADLKAVNEPE 321

RESULT 14
Q9KK27 PRELIMINARY; PRT; 869 AA.
AC Q9KK27;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Surface protein PspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]_TaxID=1313;
RP SEQUENCE FROM N.A.
RX STRAIN=G5;
RX MEDLINE=21188621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae."
RL Gene 284:63-71(2002).
DR EMBL; AF154032; AAF73801.1; -.
DR HSSP; P06653; 1H8G.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW_binding_1; 8.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGFAMS; TIGR01168; Ysirk signal; 1.
SQ SEQUENCE 869 AA; 98732 MW; AFF2B504347E0220 CRC64;
Query Match 82.7%; Score 400.5; DB 2; Length 869;
Best Local Similarity 84.8%; Pred. No. 1e-17;
Matches 84; Conservative 3; Mismatches 11; Indels 1; Gaps 1;

QY 1 LEEINSESDYAKGFRAPLOSKLDKAKKLLKLELSGKIEELDAETAELEVLKDAE 60
DB 537 LKEIDSESDYAKGFRAPLOSKLDKAKKLLKLELSGKIEELDAETAELEVLKDAE 595

QY 61 GNNVYAYFKEGKLETTAEKAELEKAEADLKAVDEPE 99
DB 596 GNNVYAYFKEGKLETTAEKAELEKAEADLKAVNEPE 634

```

```
RESULT 15
Q9LAY3 PRELIMINARY; PRT; 417 AA.
AC Q9LAY3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF10197;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071812; AAF27708.1; -.
DR HSP; P00192; 256B.
FT NON TER 417
SQ SEQUENCE 417 AA; 46960 MW; 876EAD3276506EEC CRC64;

Query Match 81.2%; Score 393; DB 2; Length 417;
Best Local Similarity 82.8%; Pred. No. 1.e-17;
Matches 82; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

Qy 1 LEEINESDSYAKGFRAPLOSKLDAKAKLLKLELSGKIELDAAIELEVLKDAE 60
Db 213 LKEIDESDSYVKEGFRAPLQSELDKQAKLSKLELSDKIDELDAEIAKLEDLKAAE 272
Qy 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVIDEPE 99
Db 273 ENNVVDYFKEGLEKTTAAKKAELKTEADLKKAVIDEPE 311
```

Search completed: June 21, 2005, 10:22:12
Job time : 62.3194 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:50:39 ; Search time 73.8459 Seconds
(without alignments)
518.502 Million cell updates/sec

Title: US-10-674-755-16

Perfect score: 485

Sequence: 1 LKEDSESDYVKGERAP.....KKAELEKAEADLKAVDEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	477	98.4	170	7 ABW02614	Abw02614 Rct135c p
2	477	98.4	181	7 ABW02596	Abw02596 0922134c
3	477	98.4	865	6 ABU08489	Abu08489 S. pneumo
4	477	98.4	929	2 AAW14593	Aaw14593 Streptoco
5	477	98.4	929	2 AAY43384	Aay43384 S. pneumo
6	477	98.4	8991	6 ABU08487	Abu08487 S. pneumo
7	474	97.7	188	2 AAW14580	Aaw14580 Streptoco
8	474	97.7	188	7 ABW02613	Abw02613 Rct129c p
9	465	95.9	588	6 ABU08491	Abu08491 Coiled co
10	465	95.9	589	2 AAY43392	Aay43392 PspC alph
11	463	95.5	204	2 AAW14578	Aaw14578 Streptoco
12	463	95.5	204	7 ABW02612	Abw02612 Rct123c p
13	462.5	95.4	180	2 AAW14562	Aaw14562 Streptoco
14	459.5	94.7	187	2 AAW14579	Aaw14579 Streptoco
15	442	91.1	1231	6 ABU08490	Abu08490 Fragment
16	440	90.7	206	2 AAW14574	Aaw14574 Streptoco
17	440	90.7	206	7 ABW02608	Abw02608 Db15c pne
18	416	85.8	198	2 AAW14581	Aaw14581 Streptoco
19	413	85.2	198	7 ABW02615	Abw02615 Rx1c pneu
20	413	85.2	315	2 AAY04375	Aay04375 Streptoco
21	413	85.2	619	2 AAR63437	Aar63437 Pneumococ
22	413	85.2	619	2 AAR87598	Aar87598 Pneumococ
23	413	85.2	619	2 AAR66911	Aar66911 Pneumococ
24	413	85.2	619	2 AAY41838	Aay41838 Streptoco
25	413	85.2	619	5 AAE18782	AAe18782 S. pneumo

26	413	85.2	619	6 ABU45778	Abu45778 Protein e
27	413	85.2	619	8 ADOS2126	Ados2126 Streptoco
28	413	85.2	648	2 AAW70336	Aaw70336 Pneumococ
29	413	85.2	648	2 AAW62274	Aaw62274 Streptoco
30	413	85.2	648	2 AAY41837	Aay41837 Streptoco
31	413	85.2	648	2 AAW87879	Aaw87879 A. pneumoc
32	413	85.2	653	2 AAW92456	Aaw92456 S. pneumoc
33	413	85.2	684	2 AAR73912	Aar73912 Streptoco
34	411	84.7	204	2 AAW14571	Aaw14571 Streptoco
35	411	84.7	204	7 ABW02605	Abw02605 Ef1019c p
36	393	81.0	653	2 AAR27150	Aar27150 PspA frag
37	390.5	80.5	289	2 AAW62276	Aaw62276 Streptoco
38	390.5	80.5	289	2 AAY41840	Aay41840 Streptoco
39	390.5	80.5	289	2 AAW87910	Aaw87910 Protein s
40	390.5	80.5	289	2 AAW92458	Aaw92458 S. pneumo
41	386	79.6	195	2 AAW14591	Aaw14591 Streptoco
42	386	79.6	195	7 ABW02625	Abw02625 Wu2c pneu
43	369	76.1	623	6 ABU08494	Abu08494 Fragment
44	365	75.3	605	6 ABU08493	Abu08493 Fragment
45	355	73.2	190	2 AAW14569	Aaw14569 Streptoco

ALIGNMENTS

RESULT 1
ABW02614
ID ABW02614 standard; protein; 170 AA.
XX
AC ABW02614;
XX
DT 12-FEB-2004 (first entry)
XX
DE Rct135c pneumococcal surface protein A (PspA) central region.
XX
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
KW immunological; gene therapy; immunostimulant.
XX
OS Unidentified.
XX
PN US6592876-B1.
XX
PD 15-JUL-2003.
XX
PF 15-SEP-1995; 95US-00529055.
XX
PR 20-APR-1993; 93US-00048896.
XX
PR 06-JUN-1995; 95US-00465746.
XX
(UABR-) UAB RES FOUND.
XX
Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
WPI; 2003-862841/80.
XX
Immunological composition for obtaining expression products used for detecting the presence of Streptococcus pneumoniae or its strain, comprises at least two different full length isolated gene encoding pneumococcal surface protein A.
XX
Example 6; SEQ ID NO 60; 121pp; English.
XX
The present invention relates to an immunological composition comprising at least 2 different full length isolated genes encoding pneumococcal surface protein A (PspAs) from different groups based on restriction fragment polymorphism analysis. The invention is useful for obtaining expression products by recombinant techniques to detect, determine, isolate or diagnose the presence of Streptococcus pneumoniae or its strain. The expression product is useful for preparing antigenic, immunological or vaccine compositions, for eliciting antibodies, an immunological response (other than or additional to antibodies) or a protective response (including antibody or other immunological response by administering compositions to a host). The invention is also useful as

CC vaccines and in gene therapy. The present sequence is Rct135c
 CC pneumococcal surface protein A (PspA) central region. This sequence is
 CC used in the exemplification of the invention

XX
 SQ Sequence 170 AA;

Query Match 98.4%; Score 477; DB 7; Length 170;
 Best Local Similarity 99.0%; Pred. No. 2.5e-36;
 Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDESDSYDKGERAPLOSQKLTGKAKLSKLELSKIDELDAEIAKLEVLKDAE 60
 DB 1 LKEIDESDSYDKGERAPLOSQKLTGKAKLSKLELSKIDELDAEIAKLEVLKDAE 60
 QY 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVIDEPE 99
 DB 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVIDEPE 99

RESULT 2
 ABW02596
 ID ABW02596 standard; protein; 181 AA.
 XX
 AC ABW02596;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 DE 0922134c pneumococcal surface protein A (PspA) central region.
 XX
 KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
 KW immunological; gene therapy; immunostimulant.
 XX
 OS Unidentified.
 XX
 PN US6592876-B1.
 XX
 PD 15-JUL-2003.
 XX
 PF 15-SEP-1995; 95US-00529055.
 XX
 PR 20-APR-1993; 93US-00048896.
 PR 06-JUN-1995; 95US-00465746.
 XX
 PA (UABR-) UAB RES FOUND.
 XX
 PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
 XX
 DR WPI; 2003-862841/80.
 XX
 PT Immunological composition for obtaining expression products used for
 PT detecting the presence of Streptococcus pneumoniae or its strain,
 PT comprises at least two different full length isolated gene encoding
 PT pneumococcal surface protein A.
 XX
 PS Example 6; SEQ ID NO 42; 121pp; English.
 XX
 CC The present invention relates to an immunological composition comprising
 CC at least 2 different full length isolated genes encoding pneumococcal
 CC surface protein A (Pspas) from different groups based on restriction
 CC fragment polymorphism analysis. The invention is useful for obtaining
 CC expression products by recombinant techniques to detect, determine,
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its
 CC strain. The expression product is useful for preparing antigenic,
 CC immunological or vaccine compositions, for eliciting antibodies, an
 CC immunological response (other than or additional to antibodies) or a
 CC protective response (including antibody or other immunological response
 CC by administering compositions to a host). The invention is also useful as
 CC vaccines and in gene therapy. The present sequence is 0922134c
 CC pneumococcal surface protein A (PspA) central region. This sequence is
 CC used in the exemplification of the invention

XX
 SQ Sequence 181 AA;

Query Match 98.4%; Score 477; DB 7; Length 181;
 Best Local Similarity 99.0%; Pred. No. 2.7e-36;
 Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDESDSYDKGERAPLOSQKLTGKAKLSKLELSKIDELDAEIAKLEVLKDAE 60
 DB 1 LKEIDESDSYDKGERAPLOSQKLTGKAKLSKLELSKIDELDAEIAKLEVLKDAE 60
 QY 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVIDEPE 99
 DB 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVIDEPE 99

RESULT 3
 ABU08489
 ID ABU08489 standard; protein; 865 AA.
 XX
 AC ABU08489;
 XX
 DT 24-JUN-2003 (first entry)
 XX
 DE S. pneumoniae pneumococcal surface protein C (PspC) protein.
 XX
 KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
 KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
 KW antibacterial.
 XX
 OS Streptococcus pneumoniae.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..37
 FT /label= signal_peptide
 FT Protein 38..865
 FT /label= Mature_PspC_protein
 XX
 PN US6500613-B1.
 XX
 PD 31-DEC-2002.
 XX
 PF 16-SEP-1996; 96US-00714741.
 XX
 PR 15-SEP-1995; 95US-00529055.
 XX
 PA (UVAL-) UNIV ALABAMA.
 XX
 PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
 PI Hollingshead S, Tart R, Brooks-Walter A;
 XX
 DR WPI; 2003-361534/34.
 DR N-PSDB; ABX95377.
 XX
 PT Isolated PspC amino acid sequence used as polymerase chain reaction or
 PT hybridization probe, comprises pneumococcal surface protein having alpha-
 PT helical, proline rich and repeat regions.
 XX
 PS Claim 3; Fig 21; 186pp; English.
 XX
 CC The present invention relates to the isolation of Streptococcus
 CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
 CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-
 CC like protein having alpha-helical, proline rich and repeat regions. The
 CC PspC and PspA proteins may be used in a vaccine to protect against
 CC pneumococcal infections. The polynucleotide sequences encoding PspC and
 CC PspA may be used for the expression of the proteins, and as PCR primers
 CC or hybridisation probes. The present sequence represents S. pneumoniae
 CC PspC protein

XX
 SQ Sequence 865 AA;

Query Match 98.4%; Score 477; DB 6; Length 865;
 Best Local Similarity 99.0%; Pred. No. 1.8e-35;
 Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKDAE 60
Db 466 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKDAE 525
QY 61 GNNNVEAYFKGLEKTTAAKKALEKAEADLKKAVIDEPE 99
Db 526 GNNNVEAYFKGLEKTTAAKKALEKAEADLKKAVIDEPE 564

RESULT 4
AAW14593
ID AAW14593 standard; protein; 929 AA.
XX
AC AAW14593;
XX
DT 17-OCT-2003 (revised)
DT 27-OCT-1997 (first entry)
XX
DE Streptococcus pneumoniae PspC.
XX
KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW bacteraemia; pneumonia.
XX
OS Streptococcus pneumoniae; strain EF6796.
XX
FH Key Location/Qualifiers
FT Peptide 1..37
FT /label= Sig_peptide
FT Protein 38..929
FT /label= Mat_protein
FT Domain 38..637
FT /label= Alpha-helix
FT Region 41..242
FT /label= Repeat_1
FT /notes= "alpha-helical repeat region"
FT Region 69..637
FT /label= Coiled-coil
FT /notes= "breaks in 7-residue periodicity of the coiled-coil occur at amino acids 136-141, 261-304 and 383-387"
FT Region 332..492
FT /label= Repeat_2
FT /notes= "alpha-helical repeat region"
FT Domain 627..689
FT /label= Proline-rich
FT Domain 913..929
FT /label= C-terminal
XX
PN WO9709994-A1.
XX
PD 20-MAR-1997.
XX
PF 16-SEP-1996; 96WO-US014819.
XX
PR 15-SEP-1995; 95US-00529055.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, McDaniel LS, Swiatlo E, Vother J, Crain MJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX
XX WPI; 1997-202002/18.
XX
XX N-PSDB; AAT61728.
XX
XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
XX in vaccines for protecting animals against S.pneumoniae infection.
XX
XX Disclosure; Fig 13; 296pp; English.
XX
XX This sequence comprises the pneumococcal protein surface C (pspC) of
XX Streptococcus pneumoniae strain EF6796. The sequence was deduced from the
XX pspC gene (AAT61728). Like PspA, PspC has an alpha-helical coiled-coil
XX region, proline-rich central region, repeat regions, with a choline
XX binding motif, and a C-terminal 17-amino acid tail. The 2 polypeptides
```

```
CC share 3 regions of high sequence identity. One is a protection-eliciting
CC region present within the alpha-helical domain. The others are the
CC proline-rich domain and a repeat domain shared with other choline-binding
CC proteins and thought to play a role in cell surface association. PspC and
CC PspA polypeptides, and their fragments, can be used in vaccines to
CC protect against S. pneumoniae infection and hence for the prevention of
CC diseases such as otitis media, meningitis, bacteraemia and pneumonia.
CC (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 929 AA;
XX
XX Query Match 98.4%; Score 477; DB 2; Length 929;
XX Best Local Similarity 99.0%; Pred. No. 1.9e-35;
XX Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKDAE 60
Db 530 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKDAE 589
QY 61 GNNNVEAYFKGLEKTTAAKKALEKAEADLKKAVIDEPE 99
Db 590 GNNNVEAYFKGLEKTTAAKKALEKAEADLKKAVIDEPE 628

RESULT 5
AAW43384
ID AAW43384 standard; protein; 929 AA.
XX
AC AAW43384;
XX
DT 27-JAN-2000 (first entry)
XX
DE S. pneumoniae PspC protein sequence.
XX
KW PspC gene; pneumococcal surface protein C; epitope; diagnosis;
KW epitopic region; immunogenic composition; vaccine composition; therapy;
KW meningitis; immunological response; otitis media; bacteraemia; pneumonia;
KW anti-PspA antibody.
XX
OS Streptococcus pneumoniae.
XX
PN WO9593940-A1.
XX
PD 28-OCT-1999.
XX
PF 23-APR-1999; 99WO-US008895.
XX
PR 23-APR-1998; 98US-0082728P.
XX
XX (UYAL-) UNIV ALABAMA.
XX
XX Briles DE, Hollingshead SK, Brooks-Walter A;
PI
XX
XX WPI; 1999-620581/53.
XX
XX N-PSDB; AAZ31956.
XX
XX New epitope polypeptides of Pneumococcal surface protein C, used to
XX develop products for immunological, immunogenic or vaccine compositions,
XX particularly against Streptococcus pneumoniae infections.
XX
XX Example; Fig 11; 109pp; English.
XX
XX This sequence is the Streptococcus pneumoniae pneumococcal surface
XX protein C. The invention relates to an isolated and/or purified
XX polypeptide comprising at least one epitope or epitopic region of
XX Pneumococcal surface protein C (PspC). The polypeptides or vectors
XX containing sequence encoding them can be used as immunogenic,
XX immunological or vaccine compositions. The compositions can be used for
XX eliciting an immunological response against Streptococcus pneumoniae
XX (SP), which can cause otitis media, meningitis, bacteraemia and
XX pneumonia. They can be used for eliciting an anti-PspA antibody. The
XX nucleic acid molecules can also be used for detecting pspC, pspA or SP
XX
```

SQ Sequence 929 AA;

Query Match 98.4%; Score 477; DB 2; Length 929;
 Best Local Similarity 99.0%; Pred. No. 1.9e-35;
 Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEYDLKEGERAPLQSKLDTKKA KLSKLELSKDIDELDAEIAKLEVLQKDAE 60
 |||||
 Db 530 LKEIDESDSEYDLKEGLRAPLQSKLDTKKA KLSKLELSKDIDELDAEIAKLEVLQKDAE 589
 |||||

Qy 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
 |||||
 Db 590 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 628
 |||||

RESULT 6

ABU08487

ID ABU08487 standard; protein; 8991 AA.

XX AC ABU08487;

XX XX

DT 24-JUN-2003 (first entry)

XX XX

DE S. pneumoniae pneumococcal surface protein A (PspA) protein.

XX XX

KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
 KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
 KW antibacterial.

XX XX

OS Streptococcus pneumoniae.

XX XX

PH Key Location/Qualifiers

FT Misc-difference 1. .8991

FT /note="All Xaa residues within this sequence are
 unknown"

FT XX

FN US6500613-B1.

XX XX

PD 31-DEC-2002.

XX XX

PF 16-SEP-1996; 96US-00714741.

XX XX

PR 15-SEP-1995; 95US-00529055.

XX XX

PA (UVAL-) UNIV ALABAMA.

XX XX

PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
 PI Hollingshead S, Tart R, Brooks-Walter A;

XX XX

DR WPI; 2003-361534/34.

XX XX

PT Isolated PspC amino acid sequence used as polymerase chain reaction or
 PT hybridization probe, comprises pneumococcal surface protein having alpha-
 PT helical, proline rich and repeat regions.

XX XX

PS Disclosure; Col 145-188; 186pp; English.

XX XX

CC The present invention relates to the isolation of Streptococcus
 CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
 CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-
 CC like protein having alpha-helical, proline rich and repeat regions. The
 CC PspC and PspA proteins may be used in a vaccine to protect against
 CC pneumococcal infections. The polynucleotide sequences encoding PspC and
 CC PspA may be used for the expression of the proteins, and as PCR primers
 CC or hybridisation probes. The present sequence represents S. pneumoniae
 CC PspA protein

XX XX

SQ Sequence 8991 AA;

Query Match 98.4%; Score 477; DB 6; Length 8991;
 Best Local Similarity 99.0%; Pred. No. 2.9e-34;
 Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEYDLKEGERAPLQSKLDTKKA KLSKLELSKDIDELDAEIAKLEVLQKDAE 60
 |||||
 Db 1 LKEIDESDSEYDLKEGLRAPLQSKLDTKKA KLSKLELSKDIDELDAEIAKLEVLQKDAE 60
 |||||

Qy 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
 |||||
 Db 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPD 99
 |||||

RESULT 7

AAW14580

ID AAW14580 standard; protein; 188 AA.

XX AC AAW14580;

XX XX

DT 17-OCT-2003 (revised)

DT 28-OCT-1997 (first entry)

XX XX

DE Streptococcus pneumoniae PspA central region.

XX XX

KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
 KW bacteraemia; pneumonia.

XX XX

OS Streptococcus pneumoniae; strain Rct135.

XX XX

PN WO9709994-A1.

XX XX

PD 20-MAR-1997.

XX XX

PF 16-SEP-1996; 96WO-US014819.

XX XX

PR 15-SEP-1995; 95US-00529055.

XX XX

PA (UABR-) UAB RES FOUND.

XX XX

PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
 PI Hollingshead S, Tart R, Brooks-Walter A;

XX XX

DR WPI; 1997-202002/18.

XX XX

PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used
 PT in vaccines for protecting animals against S.pneumoniae infection.

XX XX

PS Example 6; Fig 13; 296pp; English.

XX XX

CC This sequence shows the central portion, including the C-terminus of the
 CC alpha-helix region and some of the proline-rich region, of pneumococcal
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Rct135.
 CC Comparison of the N-terminal and central regions (AAW14533-57 and
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
 CC be used to divide the strains into several families based on sequence
 CC homologies. PspA polypeptides, or fragments of them, can be used in
 CC vaccines to protect animals against S. pneumoniae infection and hence for
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
 CC region and the immediate 5' tip of the coding sequence are likely to be
 CC the critical sequences for predicting PspA cross-reactions and vaccine
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)

XX XX

SQ Sequence 188 AA;

Query Match 97.7%; Score 474; DB 2; Length 188;
 Best Local Similarity 98.0%; Pred. No. 5.4e-36;
 Matches 97; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEYDLKEGERAPLQSKLDTKKA KLSKLELSKDIDELDAEIAKLEVLQKDAE 60
 |||||
 Db 1 LKEIDESDSEYDLKEGLRAPLQSKLDTKKA KLSKLELSKDIDELDAEIAKLEVLQKDAE 60
 |||||

Qy 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
 |||||
 Db 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPD 99
 |||||

RESULT 8
ABW02613
ID ABW02613 standard; protein; 188 AA.
XX
XX
AC ABW02613;
XX
XX
DT 12-FEB-2004 (first entry)
XX
DE Rct129c pneumococcal surface protein A (PspA) central region.
XX
XX
XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
KW immunological; gene therapy; immunostimulant.
XX
OS Unidentified.
XX
XX
PN US6592876-B1.
XX
XX
PD 15-JUL-2003.
XX
XX
PF 15-SEP-1995; 95US-00529055.
XX
XX
PR 20-APR-1993; 93US-00048896.
XX
XX
PR 06-JUN-1995; 95US-00465746.
XX
XX
PA (UABR-) UAB RES FOUND.
XX
XX
PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX
XX
XX WPI; 2003-862841/80.
XX
XX
XX Immunological composition for obtaining expression products used for
PT detecting the presence of Streptococcus pneumoniae or its strain,
PT comprises at least two different full length isolated gene encoding
PT pneumococcal surface protein A.
XX
XX
XX Example 6; SEQ ID NO 59; 121pp; English.
XX
XX
CC The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspAs) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antigenic,
CC immunological or vaccine compositions, for eliciting antibodies, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Rct129c
CC pneumococcal surface protein A (PspA) central region. This sequence is
CC used in the exemplification of the invention
XX
XX
SQ Sequence 188 AA;
Query Match 97.7%; Score 474; DB 7; Length 188;
Best Local Similarity 98.0%; Pred. No. 5.4e-36;
Matches 97; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSKDIDELDAETAKLEVQLKDAE 60
DB 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSKDIDELDAETAKLEVQLKDAE 60
QY 61 GNNVYAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 99
DB 61 GNNVYAYFKEGLEKTTAEKKALEKAEADLKKAVDEPD 99
RESULT 9
ABU08491
ID ABU08491 standard; protein; 588 AA.
XX
XX
AC ABU08491;

XX
DT 24-JUN-2003 (first entry)
XX
DE Coiled coil motif of alpha-helix of S. pneumoniae PspC protein.
XX
XX
KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
KW antibacterial.
XX
XX
OS Streptococcus pneumoniae.
XX
XX
PN US6500613-B1.
XX
XX
PD 31-DEC-2002.
XX
XX
PF 16-SEP-1996; 96US-00714741.
XX
XX
PR 15-SEP-1995; 95US-00529055.
XX
XX
PA (UYAL-) UNIV ALABAMA.
XX
XX
PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
XX
XX
XX Hollingshead S, Tart R, Brooks-Walter A;
XX
XX
XX WPI; 2003-361534/34.
XX
XX
XX Isolated PspC amino acid sequence used as polymerase chain reaction or
PT hybridization probe, comprises pneumococcal surface protein having alpha-
PT helical, proline rich and repeat regions.
XX
XX
PS Disclosure; Fig 23; 186pp; English.
XX
XX
CC The present invention relates to the isolation of Streptococcus
CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-
CC like protein having alpha-helical, proline rich and repeat regions. The
CC PspC and PspA proteins may be used in a vaccine to protect against
CC pneumococcal infections. The polynucleotide sequences encoding PspC and
CC PspA may be used for the expression of the proteins, and as PCR primers
CC or hybridization probes. The present sequence represents a coiled coil
CC motif of the alpha-helix of S. pneumoniae PspC protein
XX
XX
SQ Sequence 588 AA;
Query Match 95.9%; Score 465; DB 6; Length 588;
Best Local Similarity 99.0%; Pred. No. 1.4e-34;
Matches 96; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSKDIDELDAETAKLEVQLKDAE 60
DB 492 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSKDIDELDAETAKLEVQLKDAE 551
QY 61 GNNVYAYFKEGLEKTTAEKKALEKAEADLKKAVDE 97
DB 552 GNNVYAYFKEGLEKTTAEKKALEKAEADLKKAVDE 588
RESULT 10
AAY43392
ID AAY43392 standard; protein; 589 AA.
XX
XX
AC AAY43392;
XX
XX
DT 27-JAN-2000 (first entry)
XX
DE PspC alpha-helix coiled-coil region.
XX
XX
KW PspC gene; pneumococcal surface protein C; epitope; diagnosis;
KW epitopic region; immunogenic composition; vaccine composition; therapy;
KW meningitis; immunological response; otitis media; bacteraemia; pneumonia;
KW anti-PspA antibody.
XX
XX
OS Streptococcus pneumoniae.

```

XX PN W09953940-A1.
XX XX
XX PD 28-OCT-1999.
XX XX
XX PF 23-APR-1999; 99WO-US008895.
XX XX
XX PR 23-APR-1998; 98US-0082728P.
XX XX
XX PA (UYAL-) UNIV ALABAMA.
XX XX
XX PI Briles DE, Hollingshead SK, Brooks-Walter A;
XX XX
XX DR WPI; 1999-620581/53.
XX XX
XX PT New epitope polypeptides of Pneumococcal surface protein C, used to
XX XX
XX PF develop products for immunological, immunogenic or vaccine compositions,
XX XX
XX PT particularly against Streptococcus pneumoniae infections.
XX XX
XX PS Example 1; Fig 3; 109pp; English.
XX XX
XX CC This sequence is the coiled-coil region of the Streptococcus pneumoniae
XX XX
XX CC pneumococcal surface protein C. The invention relates to an isolated
XX XX
XX CC and/or purified polypeptide comprising at least one epitope or epitopic
XX XX
XX CC region of Pneumococcal surface protein C (PspC). The polypeptides or
XX XX
XX CC vectors containing sequence encoding them can be used as immunogenic,
XX XX
XX CC immunological or vaccine compositions. The compositions can be used for
XX XX
XX CC eliciting an immunological response against Streptococcus pneumoniae
XX XX
XX CC (SP), which can cause otitis media, meningitis, bacteraemia and
XX XX
XX CC pneumonia. They can be used for eliciting an anti-PspA antibody. The
XX XX
XX CC nucleic acid molecules can also be used for detecting pspC, pspA or SP
XX XX
XX SQ Sequence 589 AA;

Query Match 95.9%; Score 465; DB 2; Length 589;
Best Local Similarity 99.0%; Pred. No. 1.4e-34;
Matches 96; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKETDESSEDYLKEGGERAPLQSKLDTKAKLSKLELSKIDELDAEIAKLEVLQKDAE 60
Db 493 LKETDESSEDYLKEGGERAPLQSKLDTKAKLSKLELSKIDELDAEIAKLEVLQKDAE 552
Qy 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDE 97
Db 553 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDE 589

RESULT 11
AAW14578
ID AAW14578 standard; protein; 204 AA.
XX XX
XX AC AAW14578;
XX XX
XX DT 17-OCT-2003 (revised)
XX DT 28-OCT-1997 (first entry)
XX XX
XX DE Streptococcus pneumoniae PspA central region.
XX XX
XX KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
XX KW bacteraemia; pneumonia.
XX XX
XX OS Streptococcus pneumoniae; strain Rct123.
XX XX
XX PH Key Location/Qualifiers
XX FT Misc-difference 4 /note= "unidentified amino acid"
XX FT Misc-difference 8 /note= "unidentified amino acid"
XX FT
XX XX
XX PN W09709994-A1.
XX XX
XX PD 20-MAR-1997.
XX XX

```

```

PF 16-SEP-1996; 96WO-US014819.
XX XX
XX PR 15-SEP-1995; 95US-00529055.
XX XX
XX PA (UABR-) UAB RES FOUND.
XX XX
XX PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
XX PI Hollingshead S, Tart R, Brooks-Walter A;
XX XX
XX DR WPI; 1997-202002/18.
XX XX
XX PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used
XX PT in vaccines for protecting animals against S.pneumoniae infection.
XX XX
XX PS Example 6; Fig 13; 296pp; English.
XX XX
XX CC This sequence shows the central portion, including the C-terminus of the
XX CC alpha-helix region and some of the proline-rich region, of pneumococcal
XX CC surface protein A (PspA) of Streptococcus pneumoniae strain Rct123.
XX CC Comparison of the N-terminal and central regions (AAW14533-57 and
XX CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
XX CC be used to divide the strains into several families based on sequence
XX CC homologies. PspA polypeptides, or fragments of them, can be used in
XX CC vaccines to protect animals against S. pneumoniae infection and hence for
XX CC the prevention of diseases such as otitis media, meningitis, bacteraemia
XX CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
XX CC region and the immediate 5' tip of the coding sequence are likely to be
XX CC the critical sequences for predicting PspA cross-reactions and vaccine
XX CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX XX
XX SQ Sequence 204 AA;

Query Match 95.5%; Score 463; DB 2; Length 204;
Best Local Similarity 96.0%; Pred. No. 6.2e-35;
Matches 95; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYLKEGGERAPLQSKLDTKAKLSKLELSKIDELDAEIAKLEVLQKDAE 60
Db 1 IKEXDESXSSEDYLKEGLRAPLQSKLDTKAKLSKLELSKIDELDAEIAKLEVLQKDAE 60
Qy 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 12
ABW02612
ID ABW02612 standard; protein; 204 AA.
XX XX
XX AC ABW02612;
XX XX
XX DT 12-FEB-2004 (first entry)
XX XX
XX DE Rct123c pneumococcal surface protein A (PspA) central region.
XX XX
XX KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
XX KW immunological; gene therapy; immunostimulant.
XX XX
XX OS Unidentified.
XX XX
XX PH Key Location/Qualifiers
XX FT Misc-difference 1. .204 /note= "Xaa = Unknown amino acid"
XX FT
XX XX
XX PN US6592876-B1.
XX XX
XX PD 15-JUL-2003.
XX XX
XX PF 15-SEP-1995; 95US-00529055.
XX XX
XX PR 20-APR-1993; 93US-00048896.
XX PR 06-JUN-1995; 95US-00465746.
XX XX

```

PA (UABR-) UAB RES FOUND.
 XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
 XX WPI; 2003-862841/80.
 XX Immunological composition for obtaining expression products used for
 PT detecting the presence of Streptococcus pneumoniae or its strain,
 PT comprises at least two different full length isolated gene encoding
 PT pneumococcal surface protein A.
 XX Example 6; SEQ ID NO 58; 121pp; English.
 XX The present invention relates to an immunological composition comprising
 CC at least 2 different full length isolated genes encoding pneumococcal
 CC surface protein A (PspAs) from different groups based on restriction
 CC fragment polymorphism analysis. The invention is useful for obtaining
 CC expression products by recombinant techniques to detect, determine,
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its
 CC strain. The expression product is useful for preparing antigenic,
 CC immunological or vaccine compositions, for eliciting antibodies, an
 CC immunological response (other than or additional to antibodies) or a
 CC protective response (including antibody or other immunological response
 CC by administering compositions to a host). The invention is also useful as
 CC vaccines and in gene therapy. The present sequence is Rct123c
 CC pneumococcal surface protein A (PspA) central region. This sequence is
 CC used in the exemplification of the invention
 XX Sequence 204 AA;
 SQ
 Query Match 95.5%; Score 463; DB 7; Length 204;
 Best Local Similarity 96.0%; Pred. No. 6.2e-35;
 Matches 95; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSKDIDELDAEIAKLEVLKDAE 60
 Db 1 IKEXDSXSXYLKEGLRAPLQSKLDTKKAKLSKLELSKDIDELDAEIAKLEVLKDAE 60
 QY 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
 Db 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
 RESULT 13
 ID AAW14562
 XX AAW14562 standard; protein; 180 AA.
 AC AAW14562;
 XX 17-OCT-2003 (revised)
 DT 28-OCT-1997 (first entry)
 XX Streptococcus pneumoniae PspA central region.
 DE PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
 KW bacteriaemia; pneumonia.
 XX Streptococcus pneumoniae; strain 0922134c.
 OS WO9709994-A1.
 PN 20-MAR-1997.
 XX 16-SEP-1996; 96WO-US014819.
 XX 15-SEP-1995; 95US-00529055.
 XX (UABR-) UAB RES FOUND.
 XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
 PI Hollingshead S, Tart R, Brooks-Walter A;
 XX WPI; 1997-202002/18.

XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
 PT in vaccines for protecting animals against S.pneumoniae infection.
 XX Example 6; Fig 13; 296pp; English.
 XX This sequence shows the central portion, including the C-terminus of the
 CC alpha-helix region and some of the proline-rich region, of pneumococcal
 CC surface protein A (PspA) of Streptococcus pneumoniae strain 0922134c.
 CC Comparison of the N-terminal and central regions (AAW14533-57 and
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
 CC be used to divide the strains into several families based on sequence
 CC homologies. PspA polypeptides, or fragments of them, can be used in
 CC vaccines to protect animals against S. pneumoniae infection and hence for
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
 CC region and the immediate 5' tip of the coding sequence are likely to be
 CC the critical sequences for predicting PspA cross-reactions and vaccine
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)
 XX Sequence 180 AA;
 SQ
 Query Match 95.4%; Score 462.5; DB 2; Length 180;
 Best Local Similarity 98.0%; Pred. No. 5.9e-35;
 Matches 97; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSKDIDELDAEIAKLEVLKDAE 60
 Db 1 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLELSKDIDELDAEIAKLEVLKDAE 60
 QY 61 GNNVYAYFKEGLEKTTASKKAELKAEADLKKAVDEPE 99
 Db 61 GNNVYAYFKEGLEKTTASKKAELKAEADLKKAVDEPE 98
 RESULT 14
 AAW14579
 ID AAW14579 standard; protein; 187 AA.
 XX AAW14579;
 XX 17-OCT-2003 (revised)
 DT 28-OCT-1997 (first entry)
 XX Streptococcus pneumoniae PspA central region.
 DE PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
 KW bacteriaemia; pneumonia.
 XX Streptococcus pneumoniae; strain Rct129.
 OS WO9709994-A1.
 PN 20-MAR-1997.
 XX 16-SEP-1996; 96WO-US014819.
 XX 15-SEP-1995; 95US-00529055.
 XX (UABR-) UAB RES FOUND.
 XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
 PI Hollingshead S, Tart R, Brooks-Walter A;
 XX WPI; 1997-202002/18.
 XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
 PT in vaccines for protecting animals against S.pneumoniae infection.
 XX Example 6; Fig 13; 296pp; English.
 XX This sequence shows the central portion, including the C-terminus of the
 CC alpha-helix region and some of the proline-rich region, of pneumococcal

CC surface protein A (PspA) of Streptococcus pneumoniae strain Rct129.
CC Comparison of the N-terminal and central regions (AAW14533-57 and
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
CC be used to divide the strains into several families based on sequence
CC homologies. PspA polypeptides, or fragments of them, can be used in
CC vaccines to protect animals against S. pneumoniae infection and hence for
CC the prevention of diseases such as otitis media, meningitis, bacteraemia
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
CC region and the immediate 5' tip of the coding sequence are likely to be
CC the critical sequences for predicting PspA cross-reactions and vaccine
CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 187 AA;

Query Match 94.7%; Score 459.5; DB 2; Length 187;
Best Local Similarity 97.0%; Pred. No. 1.2e-34;
Matches 96; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
Qy 1 LKEIDESDSEDLKEGERAPLQSKLDTKKAKLSKLELSKDIDELDAEIAKLEVLQKDAE 60
Db 1 LKEIDESDSEDLKEGLRAPLQSKLDTKKAKLSKLELSKDIDELDAEIAKLEVLQKDAE 60
Qy 61 GNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 61 GNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPD 98

RESULT 15
ABU08490
ID ABU08490 standard; protein; 1231 AA.
XX
AC ABU08490;
XX
DT 24-JUN-2003 (first entry)
XX
DE Fragment of S. pneumoniae PspC protein.
XX
KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
KW antibacterial.
XX
OS Streptococcus pneumoniae.
XX
FN US6500613-B1.
XX
PD 31-DEC-2002.
XX
PF 16-SEP-1996; 96US-00714741.
XX
PR 15-SEP-1995; 95US-00529055.
XX
PA (UYAL-) UNIV ALABAMA.
XX
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX
DR WPI; 2003-361534/34.
XX

PT Isolated PspC amino acid sequence used as polymerase chain reaction or
PT hybridization probe, comprises pneumococcal surface protein having alpha-
PT helical, proline rich and repeat regions.
PS Disclosure; Fig 22; 186pp; English.
XX
CC The present invention relates to the isolation of Streptococcus
CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-
CC like protein having alpha-helical, proline rich and repeat regions. The
CC PspC and PspA proteins may be used in a vaccine to protect against
CC pneumococcal infections. The polynucleotide sequences encoding PspC and
CC PspA may be used for the expression of the proteins, and as PCR primers
CC or hybridisation probes. The present sequence represents a fragment of S.
CC pneumoniae PspC protein

XX
SQ Sequence 1231 AA;
Query Match 91.1%; Score 442; DB 6; Length 1231;
Best Local Similarity 65.8%; Pred. No. 4.7e-32;
Matches 98; Conservative 0; Mismatches 1; Indels 50; Gaps 1;
Qy 1 LKEIDESDSEDLKEGERAPLQSKLDTKKAKLSKLELSKDIDELDAEIAKLEVLQKDAE 54
Db 442 LKEIDESDSEDLKEGLRAPLQSKLDTKKAKLSKLELSKDIDELDAEIAKLEVLQSESE 501
Qy 55 -----QLKDAEGNNNVEAYFK 70
Db 502 DYAKEGFRAPLQSKLDAKKAELKAEADLKKAVDEPE 99
Qy 71 EGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 562 EGLEKTTAEKKAELKAEADLKKAVDEPE 590
Search completed: June 21, 2005, 10:10:15
Job time : 73.8459 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	485	100.0	99	4	US-09-147-875A-16	Sequence 16, Appl
2	477	98.4	170	4	US-08-529-055-60	Sequence 60, Appl
3	477	98.4	181	4	US-08-529-055-42	Sequence 42, Appl
4	477	98.4	854	4	US-08-714-741-40	Sequence 40, Appl
5	477	98.4	8991	4	US-08-714-741-32	Sequence 32, Appl
6	474	97.7	188	4	US-08-529-055-59	Sequence 59, Appl
7	465	95.9	141	4	US-09-286-981B-2	Sequence 2, Appl
8	465	95.9	588	4	US-08-714-741-42	Sequence 42, Appl
9	463	95.5	99	2	US-08-710-749-15	Sequence 15, Appl
10	463	95.5	204	4	US-08-529-055-58	Sequence 58, Appl
11	457	94.2	99	2	US-08-710-749-17	Sequence 17, Appl
12	456.5	94.1	100	4	US-09-147-875A-10	Sequence 10, Appl
13	448	92.4	99	4	US-09-147-875A-14	Sequence 14, Appl
14	442	91.1	1231	4	US-08-714-741-41	Sequence 41, Appl
15	440	90.7	206	4	US-08-529-055-54	Sequence 54, Appl
16	437	90.1	99	4	US-09-147-875A-15	Sequence 15, Appl
17	426	87.8	99	2	US-08-710-749-13	Sequence 13, Appl
18	423	87.2	99	2	US-08-710-749-14	Sequence 14, Appl
19	413	85.2	99	2	US-08-710-749-11	Sequence 11, Appl
20	413	85.2	198	4	US-08-529-055-61	Sequence 61, Appl
21	413	85.2	619	1	US-08-465-746-2	Sequence 2, Appl
22	413	85.2	619	1	US-08-214-164-2	Sequence 2, Appl
23	413	85.2	619	2	US-08-467-852B-3	Sequence 3, Appl
24	413	85.2	619	2	US-08-466-636-2	Sequence 2, Appl
25	413	85.2	619	2	US-08-247-491A-3	Sequence 3, Appl
26	413	85.2	619	2	US-08-319-795-2	Sequence 2, Appl
27	413	85.2	619	2	US-08-468-985-2	Sequence 2, Appl

```
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-60

Query Match 98.4%; Score 477; DB 4; Length 170;
Best Local Similarity 99.0%; Pred. No. 5.4e-38;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSDKIDELDAEIAKLEVLQKDAE 60
Db 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSDKIDELDAEIAKLEVLQKDAE 60

Qy 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDPE 99
Db 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDPE 99

RESULT 3
US-08-529-055-42
; Sequence 42, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 181 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-42

Query Match 98.4%; Score 477; DB 4; Length 181;
Best Local Similarity 99.0%; Pred. No. 5.9e-38;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSDKIDELDAEIAKLEVLQKDAE 60
Db 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSDKIDELDAEIAKLEVLQKDAE 60

Qy 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDPE 99
Db 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDPE 99

RESULT 4
US-08-714-741-40
; Sequence 40, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USSS OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 864 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-714-741-40

Query Match      98.4%; Score 477; DB 4; Length 864;
Best Local Similarity 99.0%; Pred. No. 3.6e-37;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSDKIDELDAEIAKLEVLQKDAE 60
Db 465 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLELSDKIDELDAEIAKLEVLQKDAE 524

QY 61 GNNV EAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 525 GNNV EAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 563

RESULT 5
US-08-714-741-32
; Sequence 32, Application US/08/14741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: THEREOF, EXPRESSION PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8991 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-714-741-32

Query Match      98.4%; Score 477; DB 4; Length 8991;
Best Local Similarity 99.0%; Pred. No. 5.6e-36;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSDKIDELDAEIAKLEVLQKDAE 60
Db 3667 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLELSDKIDELDAEIAKLEVLQKDAE 3726

; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-529-055-59

Query Match      97.7%; Score 474; DB 4; Length 188;
Best Local Similarity 98.0%; Pred. No. 1.2e-37;
Matches 97; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSDKIDELDAEIAKLEVLQKDAE 60
Db 1 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLELSDKIDELDAEIAKLEVLQKDAE 60

QY 61 GNNV EAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 61 GNNV EAYFKEGLEKTTAEKKAELKAEADLKKAVDEPD 99

RESULT 7
US-09-286-981B-2
; Sequence 2, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizemann, Theresa M.
; APPLICANT: Koenig, Scott
```

```
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; CURRENT FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-286-981B-2

Query Match          95.9%; Score 465; DB 4; Length 141;
Best Local Similarity 99.0%; Pred. No. 6e-37;
Matches 96; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDESDSDYLKEGERAPLQSKLDTKKA KLSKLELSKDIDELDAEIAKLEVLKDAE 60
Db 45 LKEIDESDSDYLKEGLRAPLQSKLDTKKA KLSKLELSKDIDELDAEIAKLEVLKDAE 104

Qy 61 GNNVAYFPKEGLEKTTAAKKAELKAEADLKKA VDE 97
Db 105 GNNVAYFPKEGLEKTTAAKKAELKAEADLKKA VDE 141

RESULT 8
US-08-714-741-42
; Sequence 42, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walker, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 588 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-714-741-42

Query Match          95.9%; Score 465; DB 4; Length 588;
Best Local Similarity 99.0%; Pred. No. 3.2e-36;
Matches 96; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDESDSDYLKEGERAPLQSKLDTKKA KLSKLELSKDIDELDAEIAKLEVLKDAE 60
Db 492 LKEIDESDSDYLKEGLRAPLQSKLDTKKA KLSKLELSKDIDELDAEIAKLEVLKDAE 551

Qy 61 GNNVAYFPKEGLEKTTAAKKAELKAEADLKKA VDE 97
Db 552 GNNVAYFPKEGLEKTTAAKKAELKAEADLKKA VDE 588

RESULT 9
US-08-710-749-15
; Sequence 15, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-15

Query Match          95.5%; Score 463; DB 2; Length 99;
Best Local Similarity 96.0%; Pred. No. 6.2e-37;
Matches 95; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LKEIDESDSDYLKEGERAPLQSKLDTKKA KLSKLELSKDIDELDAEIAKLEVLKDAE 60
Db 1 LKEIDESDSDYLKEGLRAPLQSKLDTKKA KLSKLELSKDIDELDAEIAKLEVLKDAE 60

Qy 61 GNNVAYFPKEGLEKTTAAKKAELKAEADLKKA VDE 99
Db 61 GNNVAYFPKEGLEKTTAAKKAELKAEADLKKA VDE 99
```

```

; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-710-749-17

Query Match          94.2%; Score 457; DB 2; Length 99;
Best Local Similarity 94.9%; Pred. No. 2.3e-36;
Matches 94; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDESSEDYLYKGERAPLOSQKLDTKAKLSKLEELSDKIDELDAETAKLEVQLKDAE 60
DB 1 LKEIDESSEDYKAGEFRAPLOSQKLDKAKKLSKLEELSDKIDELDAETAKLEVQLKDAE 60

QY 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
DB 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 12
US-09-147-875A-10
; Sequence 10, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 10
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-147-875A-10

Query Match          94.1%; Score 456.5; DB 4; Length 100;
Best Local Similarity 96.0%; Pred. No. 2.6e-36;
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 LKEIDESSEDYLYKGERAPLOSQKLDTKAKLSKLEELSDKIDELDAETAKLE-VOLKDA 59
DB 1 LKEIDESSEDYKAGEFRAPLOSQKLDKAKKLSKLEELSDKIDELDAETAKLEVQLKDA 60

QY 60 EGNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
DB 61 EGNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 100

```

```
RESULT 13
US-09-147-875A-14
; Sequence 14, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(99)
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid
US-09-147-875A-14

Query Match          92.4%; Score 448; DB 4; Length 99;
Best Local Similarity 91.9%; Pred. No. 1.6e-35;
Matches 91; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LKIDSDSDYLKEGERAPLOSKLDTKKAALKLELSKIDELDAEIAKLEVLKDAE 60
Db 1 LKIDSDSDYLKEGERAPLOSKLDTKKAALKLELSKIDELDAEIAKLEVLKDAE 60

Qy 61 GNNVVEYFKEGLEKTTAEKKAELKAEADLKKAVDPE 99
Db 61 GNNVVEYFKEGLEKTTAEKKAELKAEADLKKAVDPE 99

RESULT 14
US-08-714-741-41
; Sequence 41, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
```

```
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 1231 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: amino acid
US-08-714-741-41

Query Match          91.1%; Score 442; DB 4; Length 1231;
Best Local Similarity 65.8%; Pred. No. 1.1e-33;
Matches 98; Conservative 0; Mismatches 1; Indels 50; Gaps 1;

Qy 1 LKIDSDSDYLKEGERAPLOSKLDTKKAALKLELSKIDELDAEIAKLEVLKDAE 54
Db 442 LKIDSDSDYLKEGLRAPLOSKLDTKKAALKLELSKIDELDAEIAKLEVLQSESE 501
Qy 55 -----OLKDAEGNNNVEAYFK 70
Db 502 DYAKEGFRAPLOSKLDAKKAALKLELSKIDELDAEIAKLEDLKDAEGNNNVEAYFK 561

Qy 71 EGLEKTTAEKKAELKAEADLKKAVDPE 99
Db 562 EGLEKTTAEKKAELKAEADLKKAVDPE 590

RESULT 15
US-08-529-055-54
; Sequence 54, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 206 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

MOLECULE TYPE: peptide
US-08-529-055-54

Query Match 90.7%; Score 440; DB 4; Length 206;
Best Local Similarity 90.9%; Pred. No. 2.2e-34;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSDKIDELDAETAKLEVQLKDAE 60
Db 1 LKIDIDSDSDYAKEGLRAPLQSELDTKKAKLKLEELSGKIEELDAETIXELEVQLKDAE 60
QY 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKAVDEPE 99
Db 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKAVDEPE 99

Search completed: June 21, 2005, 10:25:21
Job time : 18.4867 secs

This Page Blank (uspto)

Result No.	Score	Query Match	Length	DB	ID	Description
1	485	100.0	99	15	US-10-674-755-16	Sequence 16, Appl
2	477	98.4	170	15	US-10-299-636-75	Sequence 75, Appl
3	477	98.4	181	15	US-10-299-636-57	Sequence 57, Appl
4	477	98.4	643	15	US-10-299-636-95	Sequence 95, Appl
5	477	98.4	670	9	US-09-748-875-63	Sequence 63, Appl
6	477	98.4	670	10	US-09-298-523B-63	Sequence 63, Appl
7	477	98.4	690	9	US-09-748-875-61	Sequence 61, Appl
8	477	98.4	690	10	US-09-298-523B-61	Sequence 61, Appl
9	477	98.4	691	9	US-09-748-875-1	Sequence 1, Appl
10	477	98.4	691	10	US-09-298-523B-1	Sequence 1, Appl
11	477	98.4	701	9	US-09-748-875-62	Sequence 62, Appl

```

US-10-299-636-75
; Sequence 75, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-75

Query Match          98.4%; Score 477; DB 15; Length 170;
Best Local Similarity 99.0%; Pred. No. 1.3e-31;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYLKEGERAPLQSKLDTKKA KLSKLELSDKIDELDAEIAKLEVLQKDAE 60
Db 1 LKEIDSDSDYLKEGERAPLQSKLDTKKA KLSKLELSDKIDELDAEIAKLEVLQKDAE 60

Qy 61 GNNNVEAYFKEGLEKTTAAEKKAELKAEADLKKA VDEPE 99
Db 61 GNNNVEAYFKEGLEKTTAAEKKAELKAEADLKKA VDEPE 99

RESULT 4
US-10-299-636-95
; Sequence 95, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 95
; LENGTH: 643
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-95

Query Match          98.4%; Score 477; DB 15; Length 643;
Best Local Similarity 99.0%; Pred. No. 5.7e-31;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYLKEGERAPLQSKLDTKKA KLSKLELSDKIDELDAEIAKLEVLQKDAE 60
Db 245 LKEIDSDSDYLKEGERAPLQSKLDTKKA KLSKLELSDKIDELDAEIAKLEVLQKDAE 304

Qy 61 GNNNVEAYFKEGLEKTTAAEKKAELKAEADLKKA VDEPE 99
Db 305 GNNNVEAYFKEGLEKTTAAEKKAELKAEADLKKA VDEPE 343

RESULT 5
US-09-748-875-63
; Sequence 63, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 670
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-63

```

```

Query Match      98.4%; Score 477; DB 9; Length 670;
Best Local Similarity 99.0%; Pred. No. 5.9e-31;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAQKLSKLEELSDKIDELDAEIAKLEVLKDAE 60
DB 498 LKEIDSDSDYLKEGLRAPLQSKLDTKKAQKLSKLEELSDKIDELDAEIAKLEVLKDAE 557

QY 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
DB 558 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 596

RESULT 6
US-09-298-523B-63
; Sequence 63, Application US/09298523B
; Publication No. US20030059438A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/298,523B
; CURRENT FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 670
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-63

Query Match      98.4%; Score 477; DB 10; Length 670;
Best Local Similarity 99.0%; Pred. No. 5.9e-31;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAQKLSKLEELSDKIDELDAEIAKLEVLKDAE 60
DB 498 LKEIDSDSDYLKEGLRAPLQSKLDTKKAQKLSKLEELSDKIDELDAEIAKLEVLKDAE 557

QY 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
DB 558 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 596

RESULT 7
US-09-748-875-61
; Sequence 61, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 61
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-61

Query Match      98.4%; Score 477; DB 9; Length 690;
Best Local Similarity 99.0%; Pred. No. 6.1e-31;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAQKLSKLEELSDKIDELDAEIAKLEVLKDAE 60
DB 529 LKEIDSDSDYLKEGLRAPLQSKLDTKKAQKLSKLEELSDKIDELDAEIAKLEVLKDAE 588

QY 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
DB 590 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 628

RESULT 10

```

```

QY 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
DB 589 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 627

RESULT 8
US-09-298-523B-61
; Sequence 61, Application US/09298523B
; Publication No. US20030059438A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/298,523B
; CURRENT FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 61
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-61

Query Match      98.4%; Score 477; DB 10; Length 690;
Best Local Similarity 99.0%; Pred. No. 6.1e-31;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAQKLSKLEELSDKIDELDAEIAKLEVLKDAE 60
DB 529 LKEIDSDSDYLKEGLRAPLQSKLDTKKAQKLSKLEELSDKIDELDAEIAKLEVLKDAE 588

QY 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
DB 589 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 627

RESULT 9
US-09-748-875-51
; Sequence 1, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-51

Query Match      98.4%; Score 477; DB 9; Length 691;
Best Local Similarity 99.0%; Pred. No. 6.1e-31;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAQKLSKLEELSDKIDELDAEIAKLEVLKDAE 60
DB 530 LKEIDSDSDYLKEGLRAPLQSKLDTKKAQKLSKLEELSDKIDELDAEIAKLEVLKDAE 589

QY 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
DB 590 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 628

RESULT 10

```

US-09-298-523B-1
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/298,523B
; CURRENT FILING DATE: 1999-04-23
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/298,523B
; CURRENT FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-1
Query Match 98.4%; Score 477; DB 10; Length 691;
Best Local Similarity 99.0%; Pred. No. 6.1e-31;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-1
Query 1 LKEIDESSEDYKGEGERAPLOSQKDTTKAKLSKLEELSDKIDELDAEIAKLEVLQKDAE 60
|||||
Db 530 LKEIDESSEDYKGEGERAPLOSQKDTTKAKLSKLEELSDKIDELDAEIAKLEVLQKDAE 589
Query 61 GNNVYAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 99
|||||
Db 590 GNNVYAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 628
RESULT 11
US-09-748-875-62
; Sequence 62, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; CURRENT FILING DATE: 2000-12-26
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 701
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-62
Query Match 98.4%; Score 477; DB 9; Length 701;
Best Local Similarity 99.0%; Pred. No. 6.2e-31;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-62
Query 1 LKEIDESSEDYKGEGERAPLOSQKDTTKAKLSKLEELSDKIDELDAEIAKLEVLQKDAE 60
|||||
Db 529 LKEIDESSEDYKGEGERAPLOSQKDTTKAKLSKLEELSDKIDELDAEIAKLEVLQKDAE 589
Query 61 GNNVYAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 99
|||||
Db 589 GNNVYAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 627
RESULT 12
US-09-298-523B-62
; Sequence 62, Application US/09298523B
; Publication No. US20030059438A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09298523B
; CURRENT FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 701
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-62
Query Match 98.4%; Score 477; DB 10; Length 701;
Best Local Similarity 99.0%; Pred. No. 6.2e-31;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-62
Query 1 LKEIDESSEDYKGEGERAPLOSQKDTTKAKLSKLEELSDKIDELDAEIAKLEVLQKDAE 60
|||||
Db 529 LKEIDESSEDYKGEGERAPLOSQKDTTKAKLSKLEELSDKIDELDAEIAKLEVLQKDAE 589
Query 61 GNNVYAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 99
|||||
Db 589 GNNVYAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 627
RESULT 13
US-09-748-875-2
; Sequence 2, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; CURRENT FILING DATE: 2000-12-26
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 701
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-2
Query Match 98.4%; Score 477; DB 9; Length 701;
Best Local Similarity 99.0%; Pred. No. 6.3e-31;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-2
Query 1 LKEIDESSEDYKGEGERAPLOSQKDTTKAKLSKLEELSDKIDELDAEIAKLEVLQKDAE 60
|||||
Db 530 LKEIDESSEDYKGEGERAPLOSQKDTTKAKLSKLEELSDKIDELDAEIAKLEVLQKDAE 589
Query 61 GNNVYAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 99
|||||
Db 590 GNNVYAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 628
RESULT 14
US-09-298-523B-2
; Sequence 2, Application US/09298523B
; Publication No. US20030059438A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/298,523B
; CURRENT FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 701
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-2
Query Match 98.4%; Score 477; DB 10; Length 701;
Best Local Similarity 99.0%; Pred. No. 6.2e-31;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-2
Query 1 LKEIDESSEDYKGEGERAPLOSQKDTTKAKLSKLEELSDKIDELDAEIAKLEVLQKDAE 60
|||||
Db 530 LKEIDESSEDYKGEGERAPLOSQKDTTKAKLSKLEELSDKIDELDAEIAKLEVLQKDAE 589
Query 61 GNNVYAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 99
|||||
Db 590 GNNVYAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 628

```
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-2

Query Match      98.4%; Score 477; DB 10; Length 707;
Best Local Similarity 99.0%; Pred. No. 6.3e-31;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKGERAPLQSKLDTKKALEKLELSDKIDELDAEIAKLEVLKDAE 60
Db 530 LKEIDSDSDYLKGERAPLQSKLDTKKALEKLELSDKIDELDAEIAKLEVLKDAE 599

QY 61 GNNVAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 99
Db 590 GNNVAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 628
```

```
RESULT 15
US-09-748-875-3
; Sequence 3, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748, 875
; PRIOR FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298, 523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-3
```

```
Query Match      98.4%; Score 477; DB 9; Length 711;
Best Local Similarity 99.0%; Pred. No. 6.3e-31;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKGERAPLQSKLDTKKALEKLELSDKIDELDAEIAKLEVLKDAE 60
Db 539 LKEIDSDSDYLKGERAPLQSKLDTKKALEKLELSDKIDELDAEIAKLEVLKDAE 598

QY 61 GNNVAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 99
Db 599 GNNVAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 637
```

Search completed: June 21, 2005, 11:18:36
Job time : 64.2388 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:53:39 ; Search time 9.9 Seconds
(without alignments)
962.168 Million cell updates/sec

Title: US-10-674-755-16

Perfect score: 485

Sequence: 1 LKEIDSDSEDLKEGERAP.....KKAELEKAEADLKKAVDEPE 99

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	413	85.2	619	2 A97887	surface protein ps
2	413	85.2	619	2 A41971	surface protein ps
3	124.5	25.7	744	2 P95013	pneumococcal surfa
4	121.5	25.1	161	2 S48396	tropomyosin TPM2 -
5	115.5	23.8	1818	1 S73852	hypothetical prote
6	110.5	22.8	852	2 D72230	conserved hypothet
7	105	21.6	1177	2 B75150	chromosome segrega
8	104	21.4	743	2 D84854	hypothetical prote
9	104	21.4	896	2 S43074	epidermal growth f
10	104	21.4	1169	2 A64505	Fill5 homolog - Met
11	103.5	21.3	1006	2 C70445	ATPase subunit of
12	103	21.2	281	2 F75216	hypothetical prote
13	103	21.2	1319	2 A28313	glued protein - fr
14	103	21.2	2116	2 A26655	myosin heavy chain
15	102.5	21.1	650	2 A11333	ABC transporter (A
16	102	21.0	395	2 AC17549	capsid protein (ba
17	102	21.0	764	2 T05409	hypothetical prote
18	101.5	20.9	229	2 S70532	outer surface prot
19	101.5	20.9	1093	2 S48460	probable membrane
20	101.5	20.9	1827	2 T16270	hypothetical prote
21	101	20.8	1179	2 F71190	probable chromosom
22	101	20.8	1312	2 T30845	probable DNA repai
23	101	20.8	2139	2 T18296	myosin heavy chain
24	100.5	20.7	233	2 S70531	bbk2.11 protein pr
25	100.5	20.7	629	2 F86351	protein T26F17.2 (
26	100.5	20.7	886	2 F69378	conserved hypothet
27	100	20.6	880	2 F75103	conserved hypothet
28	99	20.4	387	2 S57834	fcrA protein precu
29	99	20.4	388	2 A46173	Mrp4 protein - Str

30	99	20.4	1979	2 C71622	hypothetical prote
31	98.5	20.3	501	2 A44643	M protein precurs
32	98.5	20.3	990	2 H88733	protein F32E10.3 (
33	98.5	20.3	1110	2 I51116	NF-180 - sea lampr
34	98	20.2	388	2 S52536	fcrA 15 protein -
35	98	20.2	405	2 A33939	Pc gamma (Igg) rec
36	98	20.2	638	2 C64156	hypothetical prote
37	98	20.2	1078	2 T18352	protein P120 - Myc
38	98	20.2	1976	2 A59252	myosin heavy chain
39	97.5	20.1	1053	2 T51375	hypothetical prote
40	97.5	20.1	3488	2 T34418	hypothetical prote
41	97	20.0	415	2 S35760	fcrA protein precu
42	97	20.0	1133	2 T22976	hypothetical prote
43	96.5	19.9	385	2 T20410	hypothetical prote
44	96.5	19.9	445	2 T50972	probable zuotin (i
45	96.5	19.9	646	2 AH1587	bacteriophage prot

ALIGNMENTS

RESULT 1

A97887

surface protein pspA precursor [imported] - Streptococcus pneumoniae (strain R6)

C:Species: Streptococcus pneumoniae

C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004

C:Accession: A97887

R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E

r, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M

Y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A>Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A:Reference number: A97872; MUID:21429245; PMID:11544234

A:Accession: A97887

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-619 <KUR>

A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:AE007317; PIDN:AAK98925.1; PID:gl

C:Genetics:

A:Gene: pspA

Query Match 85.2%; Score 413; DB 2; Length 619;
Best Local Similarity 87.9%; Pred. No. 1.1e-20;
Matches 87; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 LKIDSDSEDLKEGERAPLQSKLDTKKAKLSKLELSKIDELDAEIAKLEVLQKDAE 60

Db 223 LKIDSDSEDLKEGERAPLQSKLDTKKAKLSKLELSKIDELDAEIAKLEVLQKDAE 282

QY 61 GNNVVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99

Db 283 ENNVVDYFKEGLEKTTAEKAELEKAEADLKKAVNEPE 321

RESULT 2

A41971

surface protein pspA precursor - Streptococcus pneumoniae

N:Alternate names: pneumococcal surface protein A

C:Species: Streptococcus pneumoniae

C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C:Accession: A41971; A60282; A33134

R:Yother, J.; Briles, D.E.

J. Bacteriol. 174, 601-609, 1992

A>Title: Structural properties and evolutionary relationships of PspA, a surface protein

A:Reference number: A41971; MUID:92105030; PMID:1729249

A:Accession: A41971

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-619 <YOT>

A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:M74122; NID:G153840; PIDN:AAA2701

A>Note: sequence extracted from NCBI backbone (NCIN:75635, NCBI:75636)

R:Talkington, D.F.; Crimmins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.


```
QY 65 ---VEAYFKEGLEKTTAEK-----KAELEKAEADLKAVDEPE 99
Db 1424 VLEIENYYKQLRLTTEKSEFDNNKNRULFEYFRKIRNEIEKKEAHIKTVLEETQ 1478

RESULT 6
D72230
conserved hypothetical protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: D72230
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: D72230
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-852 <ARN>
A:Cross-references: UNIPROT:Q9X1X1; GB:AE001806; GB:AE000512; NID:g4982196; PIDN:AAD3670
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM1636
C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 22.8%; Score 110.5; DB 2; Length 852;
Best Local Similarity 31.5%; Pred. No. 2.7;
Matches 28; Conservative 20; Mismatches 30; Indels 11; Gaps 2;

QY 6 ESDSDYLKEGRAPLQSKLDTKKAKLSKLE-----LSDKIDELDAETAKLEVQLKDAEG 61
Db 506 EKIEBELHRLGYSEDLLEKDKRKLKIEERHSISQKITAADVQISQIENQLKEIKG 565

QY 62 NNNVEAYFKEGLEKTTAEKAELEKAEAD 90
Db 566 E-----IEAKRETLEQREEMDLQKSD 587

RESULT 7
B75150
chromosome segregation protein (smc1) PAB2109 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: B75150
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A:Reference number: A75001
A:Accession: B75150
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1177 <KAW>
A:Cross-references: UNIPROT:Q9V1R8; GB:AJ248284; GB:AL096836; NID:g5457730; PIDN:CAB4928
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB2109
C:Superfamily: chromosome segregation protein SMC1

Query Match 21.6%; Score 105; DB 2; Length 1177;
Best Local Similarity 31.1%; Pred. No. 8.8;
Matches 33; Conservative 26; Mismatches 35; Indels 12; Gaps 4;

QY 2 KEIDSDSDYLKEGRAPLQSKLDTKKAKL-----SKLELSDKIDELDAETAKLEVQLK 57
Db 403 KSLYENADIKRLAEKRLSRITTLKAKLPQIREVEVKRLKLEEKKAELSNNVENKIS 462

QY 58 D-AEGNNNVAYFKEGLEKTTAEK---AELEKAEADLKAVDEPE 99
Db 463 SISQRRKVE-----EELEKKTSELOKVSSELSLESLRELKAEQAQSE 504
```

```
RESULT 8
D84854
hypothetical protein At2g42480 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: D84854
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84854
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-743 <STO>
A:Cross-references: UNIPROT:Q9SLB2; GB:AE002093; NID:g4567318; PIDN:AAD23729.1; GSPDB:GN
C:Genetics:
A:Gene: At2g42480
A:Map position: 2

Query Match 21.4%; Score 104; DB 2; Length 743;
Best Local Similarity 34.9%; Pred. No. 6.5;
Matches 30; Conservative 16; Mismatches 30; Indels 10; Gaps 3;

QY 6 ESDSDYLKEGRAPLQSKLDTKKAKLSKLE-ELSDKIDELDAETAKLEVQLKDAEGNNN 64
Db 260 EQDIEERLKNLEGMFEFDSKLSKLDLSKLDLISLERKKSVDADGSRVQQLSERVKD----- 313

QY 65 VEAYFKEGLEKTTAEKAELEKAEAD 90
Db 314 IELILKSKLEEVSSSEKK---KKADAD 336

RESULT 9
S43074
epidermal growth factor receptor substrate - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S43074; I38525
R:Bernard, O.A.; Mauchauffe, M.; Mecucci, C.; van den Berghe, H.; Berger, R.
Oncogene 9, 1039-1045, 1994
A:Title: A novel gene, AF-1p, fused to HRX in t(1;11)(p32;q23), is not related to AF-4, i
A:Reference number: S43074; MUID:94181254; PMID:8134107
A:Accession: S43074
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-896 <BER>
A:Cross-references: UNIPROT:P42566; EMBL:Z29064; NID:g470034; PIDN:CAA82305.1; PID:g47003
R:Wong, W.T.; Kraus, M.H.; Carlomagno, F.; Zelano, A.; Druck, T.; Croce, C.M.; Huebner, J
Oncogene 9, 1591-1597, 1994
A:Title: The human eps15 gene, encoding a tyrosine kinase substrate, is conserved in evol
A:Reference number: I38525; MUID:94239734; PMID:8183552
A:Accession: I38525
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-821,'M',823-896 <RES>
A:Cross-references: EMBL:U07707; NID:g466259; PIDN:AAA52101.1; PID:g466260
C:Genetics:
A:Gene: GDB:EPS15; AF-1P; MLT5
A:Cross-references: GDB:360337; OMIM:600051
A:Map position: lp32-1p32

Query Match 21.4%; Score 104; DB 2; Length 896;
Best Local Similarity 28.7%; Pred. No. 7.8;
Matches 29; Conservative 22; Mismatches 44; Indels 6; Gaps 2;

QY 3 EIDESDSDYLKE--GERAPLQSKLDTKKAKLSKL-----EELSCKIDELDAETAKLEVQL 56
Db 353 EQDLKEKEDTIKQRTSEVDLQDQVRENTNQLQAQKQVQQLDELDELDEQKAQLEQL 412

QY 57 KDAEGNNNVAYFKEGLEKTTAEKAELEKAEADLKAVDE 97
```

Db 413 KEVRKKCAEAQILSSILKABLTQSQISTYEBELAKAREE 453

RESULT 10

A64505

P115 homolog - Methanococcus jannaschii

C;Species: Methanococcus jannaschii

C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 02-Jun-2000

C;Accession: A64505

R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinscock, K.G.; Merrick, J.M.; Glodek, A.; rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A. Science 273, 1058-1073, 1996

A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.A.; Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A;Reference number: A64300; MUID:96337999; PMID:8688087

A;Accession: A64505

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-1169 <BUL>

A;Cross-references: GB:U67604; GB:L77117; NID:gl592224; PID:gl500543; TIGR:MJ1643

C;Genetics:

A;Map position: FOR1623481-1626990

C;Superfamily: chromosome segregation protein SMC1

Query Match 21.4%; Score 104; DB 2; Length 1169;

Best Local Similarity 27.1%; Pred. No. 10;

Matches 29; Conservative 27; Mismatches 33; Indels 18; Gaps 3;

Qy 1 LKSIDSDSDYLKGERAPLQSKLD-----TKKAKLSKLEELSDKIDELDAEIAKLEVQ 55

Db 799 LKRWNETEGELKILEKAKLKNKIDKGLTLVKEILPKTEELNKKVSELINKKVLKLN 858

Qy 56 LKDAEGNNVEAFKGELEKTTA---EKKALEKAEADLKAVDEPE 99

Db 859 I-----SPYKESIEKNLSILEEKRYRYBELAKNLKELTEKKE 895

RESULT 11

C70445

ATPase subunit of ATP-dependent proteinase (EC 3.4.-.-) - Aquifex aeolicus

C;Species: Aquifex aeolicus

C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004

C;Accession: C70445

R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V

Nature 392, 353-358, 1998

A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A;Reference number: A70300; MUID:98196666; PMID:9537320

A;Accession: C70445

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-1006 <AQF>

A;Cross-references: UNIPROT:O67588; GB:AE000750; NID:g2983999; PIDN:AAC07550.1; PID:g298

A;Experimental source: strain VF5

C;Genetics:

A;Gene: clpB

C;Superfamily: endopeptidase Clp ATP-binding chain

C;Keywords: hydrolase

Query Match 21.3%; Score 103.5; DB 2; Length 1006;

Best Local Similarity 32.7%; Pred. No. 9.5;

Matches 35; Conservative 22; Mismatches 27; Indels 23; Gaps 5;

Qy 1 LKSIDSDSDYLKGERAPLQSKLDTKKAKLSK-LEELSDKIDELDAEIAKLEV 54

Db 552 IKALEEQIIEANLKG DYKE-----AQLKIEKAKLEKQELLGKVGVEAKIAELKK 604

Qy 55 QLKDAEGNNNVEAFKGELEKTTAETKKALE-----KAEADLKAVDE 97

Db 605 KIEE-----LDEKIKAEASKGDYKEAEKAKLEKELKKLEQE 645

RESULT 12

F75216

hypothetical protein PAB2181 - Pyrococcus abyssi (strain Orsay)

C;Species: Pyrococcus abyssi

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C;Accession: F75216

R;anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru

A;Reference number: A75001

A;Accession: F75216

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-281 <RAW>

A;Cross-references: UNIPROT:Q9V217; GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB49181

A;Experimental source: strain Orsay

C;Genetics:

A;Gene: PAB2181

Query Match 21.2%; Score 103; DB 2; Length 281;

Best Local Similarity 27.6%; Pred. No. 2.9;

Matches 34; Conservative 29; Mismatches 32; Indels 28; Gaps 5;

Qy 1 LKE-IDSDSDYLKGERAPLQSKLDTKKAKL-----SKLEELSDKIDELDA 47

Db 120 IKEVWAREEYKLLKEVEK-LKQEFEEVKAKTEAAAELESLEKAKKEIEELKGKVEKLEQ 177

Qy 48 EIAKLEVLQKDAEGN-----NNVEAYFKEGLEKTTAKKAELEKAEADLKAVD 96

Db 178 EKKELEKLLKSESVKLMVEYAKAKRAELEAKLREYEEKVKREE--ELERKVSLEERSLN 235

Qy 97 EPE 99

Db 236 EYE 238

RESULT 13

A28313

glued protein - fruit fly (Drosophila melanogaster)

C;Species: Drosophila melanogaster

C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004

C;Accession: A28313

R;Swaroop, A.; Swaroop, M.; Garen, A. Proc. Natl. Acad. Sci. U.S.A. 84, 6501-6505, 1987

A;Title: Sequence analysis of the complete cDNA and encoded polypeptide for the glued pr

A;Reference number: A28313; MUID:87317680; PMID:2819881

A;Accession: A28313

A;Molecule type: DNA; mRNA

A;Residues: 1-1319 <SWA>

A;Cross-references: UNIPROT:PI3496

A;Note: the authors' translation is inconsistent with the nucleotide sequence in the reg

C;Genetics:

A;Gene: FlyBase:Gl

A;Cross-references: FlyBase:FBgn0001108

A;Introns: 18/2; 475/3

C;Keywords: cytoskeleton; glycoprotein

F;397,590,771,888,980,1110,1127,1133,1142/Binding site: carbohydrate (Asn) (covalent) #st

Query Match 21.2%; Score 103; DB 2; Length 1319;

Best Local Similarity 32.7%; Pred. No. 13;

Matches 33; Conservative 20; Mismatches 26; Indels 22; Gaps 4;

Qy 1 LKSIDSDSDYLKGERAPLQSKLDTKKAKLSK-----EELSDKIDELDAEIAKLEVQ 56

Db 429 LRDLASHDKHDIQK-----LSKELEMKRSEVTELETKELSAKIDELEAIADLQEQV 482

Qy 57 KDAEGNNNVEAYFKEGLEKTTAKKAELEKAEADLKAVDE 97

Db 483 DAALG-----AEEMVEQLAEKKMELE-----DKVKLLDEE 511

RESULT 14

A26655

myosin heavy chain [similarity] - slime mold (Dictyostelium discoideum)
N:Contains: myosin Arpase (EC 3.6.4.1)
C:Species: Dictyostelium discoideum
C>Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 09-Jul-2004
C:Accession: A26655; A24728; S00250
R:Warrick, H.M.; De Lozanne, A.; Leinwand, L.A.; Spudich, J.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 9433-9437, 1986
A:Title: Conserved protein domains in a myosin heavy chain gene from Dictyostelium discoideum
A:Reference number: A26655; MUID:87092266; PMID:3540939
A:Accession: A26655
A:Molecule type: DNA
A:Residues: 1-2116 <WAR>
A:Cross-references: UNIPROT:P08799; GB:M14628; GB:M11938; NID:G167834; PIDN:AAA33227.1;
R:DeLozanne, A.; Lewis, M.; Spudich, J.A.; Leinwand, L.A.
Proc. Natl. Acad. Sci. U.S.A. 82, 6807-6810, 1985
A:Reference number: A24728; MUID:86016788; PMID:3901008
A:Accession: A24728
A:Molecule type: mRNA
A:Residues: 2035-2116
R:Wagie, G.; Noegel, A.; Scheel, J.; Gerisch, G.
FEBS Lett. 227, 71-75, 1988
A:Title: Phosphorylation of threonine residues on cloned fragments of the Dictyostelium
A:Reference number: S00250; MUID:88112256; PMID:2828113
A:Accession: S00250
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1734-1893 <WAG>
C:Comment: The rod domain is highly periodic, containing a pattern of 7-residue repeats
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; ATP; coiled coil; hydrolase; nucleotide binding; P-loop; phosphate binding site
F:1-818/Domain: globular head <HED>
F:89-747/Domain: myosin motor domain homology <MMOT>
F:179-186/Region: nucleotide-binding motif A (P-loop)
F:819-2116/Domain: alpha-helical rod <ROD>

Query Match 21.2%; Score 103; DB 2; Length 2116;
Best Local Similarity 29.2%; Pred. No. 21;
Matches 38; Conservative 17; Mismatches 43; Indels 32; Gaps 5;

QY 1 LKEIDSDSE--DYLKEGERAPLQSKLDTKKAKLSKLELSKID-----E 44
DB 1453 IKRLNEELSELRLSVLEADER-CNSAIKAKKTAESLSKDEIDANNNAKAKAEKSK 1511
QY 45 LDAETIAKLVQLKDAEGNNVVEAYFKEG-----LEKTTA-----EKKAELKAEA 89
DB 1512 LEVRVAELSESLDGSQTVNVEFIRKDAEIDDLRLDRETESRIKSDKDKNTRKQFA 1571
QY 90 DLKKAVIDEPE 99
DB 1572 DLEAKVEEAQ 1581

RESULT 15
AII333
ABC transporter (ATP-binding protein) homolog lmo2073 [imported] - Listeria monocytogenes
C:Species: Listeria monocytogenes
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AII333
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mak, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AII333
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-650 <GLA>
A:Cross-references: UNIPROT:Q8Y519; GB:NC_003210; PIDN:CAD00151.1; PID:G16411543; GSPDB:
C:Genetics:

A:Gene: lmo2073
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 21.1%; Score 102.5; DB 2; Length 650;
Best Local Similarity 31.8%; Pred. No. 7.2;
Matches 35; Conservative 22; Mismatches 32; Indels 21; Gaps 4;

QY 2 KEIDSDSDYLKEGERAPLQS---KLDTKKAK-----LSKLELSKIDELDAEI 49
DB 541 KELARLDAEDRRKGEQVEATASVRKLNQEEKEQQLRQRKRKLEETKSEETDEKI 600
QY 50 AKLEVQLKDAEGNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVIDEPE 99
DB 601 AELEQLTNPE-----VFQDHEKALEIT-----QELDAVRAKDGKELMEWE 641

Search completed: June 21, 2005, 10:12:02
Job time : 10.9 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:54:24 ; Search time 61.3194 Seconds
(without alignments)
826.751 Million cell updates/sec

Title: US-10-674-755-16
Perfect score: 485
Sequence: 1 LKEIDESDSEYVKEGERAP.....KKAELEKADLKKAVDEPE 99

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 03: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	477	98.4	739	2	Q9RQT4	Q9RQT4 streptococc
2	477	98.4	820	2	Q9RQT1	Q9RQT1 streptococc
3	477	98.4	929	2	Q9KK19	Q9KK19 streptococc
4	477	98.4	929	2	Q9ZAY5	Q9ZAY5 streptococc
5	458	94.4	437	2	Q9LAY4	Q9LAY4 streptococc
6	450	92.8	249	2	Q9L575	Q9L575 streptococc
7	444	91.5	99	2	Q8KK4	Q8KK4 streptococc
8	441	90.9	224	2	Q8GNS8	Q8GNS8 streptococc
9	441	90.9	426	2	Q9LAY5	Q9LAY5 streptococc
10	437	90.1	395	2	Q9LAY2	Q9LAY2 streptococc
11	437	90.1	408	2	Q9LAY0	Q9LAY0 streptococc
12	433.5	89.4	869	2	Q9KK27	Q9KK27 streptococc
13	413	85.2	619	2	Q54972	Q54972 streptococc
14	413	85.2	619	2	Q8DR10	Q8DR10 streptococc
15	411	84.7	417	2	Q9LAY3	Q9LAY3 streptococc
16	391	80.6	415	2	Q9LAY1	Q9LAY1 streptococc
17	330.5	68.1	222	2	Q9L577	Q9L577 streptococc
18	330.5	68.1	225	2	Q9L591	Q9L591 streptococc
19	330.5	68.1	262	2	Q9L576	Q9L576 streptococc
20	330.5	68.1	415	2	Q9LAY7	Q9LAY7 streptococc
21	329.5	67.9	394	2	Q9LAY6	Q9LAY6 streptococc
22	329.5	67.9	395	2	Q9LAZ1	Q9LAZ1 streptococc
23	327.5	67.5	194	2	Q9L5B5	Q9L5B5 streptococc
24	327.5	67.5	218	2	Q6UBB2	Q6UBB2 streptococc
25	327.5	67.5	233	2	Q9L568	Q9L568 streptococc
26	327.5	67.5	236	2	Q9L569	Q9L569 streptococc
27	327.5	67.5	243	2	Q9L564	Q9L564 streptococc
28	327.5	67.5	243	2	Q9L567	Q9L567 streptococc
29	327.5	67.5	244	2	Q9L565	Q9L565 streptococc
30	327.5	67.5	246	2	Q9L578	Q9L578 streptococc
31	327.5	67.5	247	2	Q9L566	Q9L566 streptococc

32	327.5	67.5	249	2	Q9L570	Q9L570 streptococc
33	327.5	67.5	254	2	Q9L563	Q9L563 streptococc
34	327.5	67.5	401	2	Q9LAZ2	Q9LAZ2 streptococc
35	326.5	67.3	255	2	Q9L581	Q9L581 streptococc
36	326.5	67.3	255	2	Q9L5B6	Q9L5B6 streptococc
37	323.5	66.7	416	2	Q9LAY8	Q9LAY8 streptococc
38	319.5	65.9	393	2	Q9LAZ3	Q9LAZ3 streptococc
39	316.5	65.3	406	2	Q9LAZ0	Q9LAZ0 streptococc
40	315.5	65.1	340	2	Q8KK5	Q8KK5 streptococc
41	310.5	64.0	237	2	Q9L592	Q9L592 streptococc
42	310.5	64.0	395	2	Q9LAY9	Q9LAY9 streptococc
43	301.5	62.2	207	2	O8GNS9	O8GNS9 streptococc
44	195.5	40.3	653	2	O34097	O34097 streptococc
45	174.5	36.0	246	2	Q9L5B4	Q9L5B4 streptococc

ALIGNMENTS

RESULT 1
Q9RQT4 ID Q9RQT4 PRELIMINARY; PRT; 739 AA.
AC Q9RQT4;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Hypothetical protein pspC (fragment).
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El34;
RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic protein, PspC, which elicits cross-reactive antibodies to PspA and provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; AF068647; AAF13457.1;
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos YSIRK.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW binding_1; 1.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; YSIRK_signal; 1.
DR TIGRFAMs; TIGR01168; YSIRK_signal; 1.
KW Hypothetical protein.
FT NON TER 739
SQ SEQUENCE 739 AA; 33960 MW; 7EE2F2F676ABF989 CRC64;

Query Match 98.4%; Score 477; DB 2; Length 739;
Best Local Similarity 99.0%; Pred. No. 2.4e-22;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDESDSEYVKEGERAPLQSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKDAE 60
DB 537 LKEIDESDSEYVKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKDAE 596
QY 61 GNNVVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99
DB 597 GNNVVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 635

RESULT 2
Q9RQT1 ID Q9RQT1 PRELIMINARY; PRT; 820 AA.
AC Q9RQT1;
DT 01-MAY-2000 (Tremblrel. 13, Created)

```

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein pspC (Fragment).
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9163;
RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, pspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; AF068650; AAF13460.1; -.
DR HSP; P04268; IIC2.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW binding_1; 1.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRfams; TIGR01168; Ysirk_signal; 1.
KW Hypothetical protein.
FT NON_TPR 820
SQ SEQUENCE 820 AA; 91752 MW; 33C095849ABB0942 CRC64;

Query Match 98.4%; Score 477; DB 2; Length 820;
Best Local Similarity 99.0%; Pred. No. 2.6e-22;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEDLKGEGRAPLQSKLDTKKAQKLELSKIDELDAEIAKLEVLQKDAE 60
Db 530 LKEIDESDSEDLKGEGRAPLQSKLDTKKAQKLELSKIDELDAEIAKLEVLQKDAE 589

Qy 61 GNNVVEAYFKEGLEKTTAAEKKAELKAEADLKKAVDPE 99
Db 590 GNNVVEAYFKEGLEKTTAAEKKAELKAEADLKKAVDPE 628

RESULT 3
Q9KK19 ID Q9KK19 PRELIMINARY; PRT; 929 AA.
AC Q9KK19
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein PspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=srcf10;
RX MEDLINE=21888621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli P., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae.";
RL Gene 284:63-71(2002).
DR EMBL; AF154037; AAF73809.1; -.
DR HSP; P06653; IH8G.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.

```

```

DR Pfam; PF01473; CW binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRfams; TIGR01168; Ysirk_signal; 1.
SQ SEQUENCE 929 AA; 105003 MW; 2DC82933302FAFA64 CRC64;

Query Match 98.4%; Score 477; DB 2; Length 929;
Best Local Similarity 99.0%; Pred. No. 2.9e-22;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEDLKGEGRAPLQSKLDTKKAQKLELSKIDELDAEIAKLEVLQKDAE 60
Db 530 LKEIDESDSEDLKGEGRAPLQSKLDTKKAQKLELSKIDELDAEIAKLEVLQKDAE 589

Qy 61 GNNVVEAYFKEGLEKTTAAEKKAELKAEADLKKAVDPE 99
Db 590 GNNVVEAYFKEGLEKTTAAEKKAELKAEADLKKAVDPE 628

RESULT 4
Q9ZAY5 ID Q9ZAY5 PRELIMINARY; PRT; 929 AA.
AC Q9ZAY5
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein C.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF6796;
RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, pspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; U72655; AAD00184.1; -.
DR HSP; P06653; IHXC.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRfams; TIGR01168; Ysirk_signal; 1.
SQ SEQUENCE 929 AA; 104991 MW; 2DC82933302FFB081 CRC64;

Query Match 98.4%; Score 477; DB 2; Length 929;
Best Local Similarity 99.0%; Pred. No. 2.9e-22;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEDLKGEGRAPLQSKLDTKKAQKLELSKIDELDAEIAKLEVLQKDAE 60
Db 530 LKEIDESDSEDLKGEGRAPLQSKLDTKKAQKLELSKIDELDAEIAKLEVLQKDAE 589

Qy 61 GNNVVEAYFKEGLEKTTAAEKKAELKAEADLKKAVDPE 99
Db 590 GNNVVEAYFKEGLEKTTAAEKKAELKAEADLKKAVDPE 628

RESULT 5
Q9LAY4 ID Q9LAY4 PRELIMINARY; PRT; 437 AA.
AC Q9LAY4
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

```

Qy 1 LKEIDSDSEDYLKEGERAPIQSKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLQKDAE 60

```

DR InterPro; IPR009082; his_kin_homodim.
FT NON TER 1 224
FT NON TER 224 224
SQ SEQUENCE 224 AA; 23418 MW; 48674E27AFB66A95 CRC64;

Query Match 90.9%; Score 441; DB 2; Length 224;
Best Local Similarity 90.9%; Pred. No. 1.5e-20;
Matches 90; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYLKEGRAPLQSKLDTKAKLSKLELSKIDELDAEIAKLEVLKDAE 60
Db 17 LKIDNESDSDYVKEGRAPLQSELDTKAKLLKLELSKIEELDAEIAEVLKDAE 76
Qy 61 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 99
Db 77 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 115

RESULT 9
Q9LAY5 PRELIMINARY; PRT; 426 AA.
AC Q9LAY5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSB5;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071810; AAF27706.1; -.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS: PR00194; TROPOMYOSIN.
FT NON TER 426
SQ SEQUENCE 426 AA; 46534 MW; 81AA11348CBE6634 CRC64;

Query Match 90.9%; Score 441; DB 2; Length 426;
Best Local Similarity 90.9%; Pred. No. 2.7e-20;
Matches 90; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYLKEGRAPLQSKLDTKAKLSKLELSKIDELDAEIAKLEVLKDAE 60
Db 215 LKIDNESDSDYVKEGRAPLQSELDTKAKLLKLELSKIEELDAEIAEVLKDAE 274
Qy 61 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 99
Db 275 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 313

RESULT 10
Q9LAY2 PRELIMINARY; PRT; 395 AA.
AC Q9LAY2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX Streptococcus.

OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF6796;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071813; AAF27709.1; -.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR PRINTS: PR000533; TROPOMYOSIN.
FT NON TER 395
SQ SEQUENCE 395 AA; 42963 MW; 58E6EF956BCBCC1E CRC64;

Query Match 90.1%; Score 437; DB 2; Length 395;
Best Local Similarity 90.9%; Pred. No. 4.5e-20;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYLKEGRAPLQSKLDTKAKLSKLELSKIDELDAEIAKLEVLKDAE 60
Db 225 LKEINESDSDYAKEGFRAPLQSKLDAKAKLLKLELSKIEELDAEIAEVLKDAE 284
Qy 61 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 99
Db 285 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 323

RESULT 11
Q9LAY0 PRELIMINARY; PRT; 408 AA.
AC Q9LAY0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9163;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071815; AAF27711.1; -.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin_alc_DH_like.
FT NON TER 408
SQ SEQUENCE 408 AA; 44254 MW; 4F64D874217297EF CRC64;

Query Match 90.1%; Score 437; DB 2; Length 408;
Best Local Similarity 90.9%; Pred. No. 4.6e-20;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYLKEGRAPLQSKLDTKAKLSKLELSKIDELDAEIAKLEVLKDAE 60
Db 228 LKEINESDSDYAKEGFRAPLQSKLDAKAKLLKLELSKIEELDAEIAEVLKDAE 287
Qy 61 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 99
Db 288 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 326

RESULT 12

```



```
Q9KK27
ID Q9KK27 PRELIMINARY; PRT; 869 AA.
AC Q9KK27;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Surface protein PspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=95;
RX MEDLINE=21888621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae.";
RL Gene 284:63-71(2002).
DR EMBL; AF154032; AAF73801.1; -.
DR HSSP; P06653; 1H8G.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR Pfam; PF01473; CW_binding_1; 8.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
SQ SEQUENCE 869 AA; 98732 MW; AFF2E504347E0220 CRC64;

Query Match 89.4%; Score 433.5; DB 2; Length 869;
Best Local Similarity 91.9%; Pred. No. 1.5e-19;
Matches 91; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSKIDELDAETAKLEVQLKDAE 60
Db 537 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSKIDELDVN-CNLRSQLKDAE 595

QY 61 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 99
Db 596 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 634

RESULT 13
Q54972 PRELIMINARY; PRT; 619 AA.
ID Q54972;
AC Q54972;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pneumococcal surface protein A precursor.
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92105030; PubMed=1729249;
RA Yother J., Briles D.E.;
RT "Structural properties and evolutionary relationships of PspA, a
RT surface protein of Streptococcus pneumoniae, as revealed by sequence
RT analysis.";
RL J. Bacteriol. 174:601-609(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Yother J., Briles D.E.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; W74122; AAA27018.1; -.
DR PIR; A41971; A41971.
DR PIR; A97887; A97887.
```

```
RESULT 15
Q9LAY3 PRELIMINARY; PRT; 417 AA.
AC Q9LAY3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ef10197;
RX MEDLINE=2048953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K.; Becker R.; Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
  in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071812; AAF27708.1; -.
DR HSSP; P00192; 256B.
FT NON TER 417
SQ SEQUENCE 417 AA; 46960 MW; 876EAD3276506EEC CRC64;

Query Match      84.7%; Score 411; DB 2; Length 417;
Best Local Similarity 86.9%; Pred. No. 2e-18;
Matches 86; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEDLKXGGRAPLQSKLDTKAKLSKLEELSDKIDELDAEIAKLEVLKDAE 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 213 LKEIDESDSEDLKXGGRAPLQSKLDTKAKLSKLEELSDKIDELDAEIAKLEVLKDAE 272
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVIDEPE 99
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 273 ENNVVDYFKEGLEKTTAEKKAELKAEADLKKAVIDEPE 311
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

Search completed: June 21, 2005, 10:22:12
Job time : 61.3194 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:50:39 ; Search time 73.8459 Seconds
(without alignments)
518.502 Million cell updates/sec

Title: US-10-674-755-17

Perfect score: 465

Sequence: 1 PKRINLSQVXLKXVCRAP.....KKAELXAXADLKKAVIDEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	453	97.4	190	2	AAW14569 Streptococ
2	453	97.4	193	7	ABW02603 Bg9163c p
3	453	97.4	8991	6	ABU08487 S. pneumo
4	361	77.6	206	2	AAW14574 Streptoco
5	361	77.6	206	7	ABW02608 Db15c pne
6	355	76.3	170	7	ABW02614 Rct135c p
7	355	76.3	181	7	ABW02596 0922134c
8	355	76.3	865	6	ABU08489 S. pneumo
9	355	76.3	929	2	AAW14593 Streptoco
10	355	76.3	929	2	AAW14593 Streptoco
11	354	76.1	204	2	AAW14578 Streptoco
12	354	76.1	204	7	ABW02612 Rct132c p
13	352	75.7	188	2	AAW14580 Streptoco
14	352	75.7	188	7	ABW02613 Rct129c p
15	347	74.6	1231	6	ABU08490 Fragment
16	343	73.8	588	6	ABU08491 Coiled co
17	343	73.8	589	2	AAW14392 PspC alph
18	340.5	73.2	180	2	AAW14562 Streptoco
19	337.5	72.6	187	2	AAW14579 Streptoco
20	312	67.1	605	6	ABU08493 Fragment
21	307	66.0	198	2	AAW14581 Streptoco
22	304	65.4	198	7	ABW02615 Rx1c pneu
23	304	65.4	315	2	AAW04375 Streptoco
24	304	65.4	619	2	AAR63437 Pneumococ
25	304	65.4	619	2	AAR87598 Pneumococ

26	304	65.4	619	2	AAR86911 Pneumococ
27	304	65.4	619	2	AAW14838 Streptoco
28	304	65.4	619	5	AAE18782 S. pneumo
29	304	65.4	619	6	ABU45778 Protein e
30	304	65.4	619	8	ADO52126 Streptoco
31	304	65.4	648	2	AAW70336 Pneumococ
32	304	65.4	648	2	AAW62274 Streptoco
33	304	65.4	648	2	AAW14837 Streptoco
34	304	65.4	648	2	AAW87879 A. pneumoc
35	304	65.4	653	2	AAW92456 S. pneumo
36	304	65.4	684	2	AAW73912 Streptoco
37	297	63.9	204	2	AAW14571 Streptoco
38	297	63.9	204	7	ABW02605 Bf1018c p
39	284	63.2	289	2	AAW62276 Streptoco
40	294	63.2	289	2	AAW14840 Streptoco
41	294	63.2	289	2	AAW87910 Protein s
42	294	63.2	289	2	AAW92458 S. pneumo
43	293	63.0	653	2	AAW27150 PspA frag
44	277	59.6	195	2	AAW14591 Streptoco
45	277	59.6	195	7	ABW02625 Wuzc pneu

ALIGNMENTS

RESULT 1

AAW14569

ID AAW14569 standard; protein; 190 AA.

XX AAW14569;

XX AC AAW14569;

DT 17-OCT-2003 (revised)

DT 28-OCT-1997 (first entry)

XX

DE Streptococcus pneumoniae PspA central region.

DE

KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis; bacteraemia; pneumonia.

KW

XX Streptococcus pneumoniae; strain Bg8743.

XX

OS

FH Key Location/Qualifiers

FT Misc-difference 21 /note= "unidentified amino acid"

FT Misc-difference 24 /note= "unidentified amino acid"

FT Misc-difference 95 /note= "unidentified amino acid"

FT Misc-difference 97 /note= "unidentified amino acid"

FT Misc-difference 186 /note= "unidentified amino acid"

XX WO9709994-A1.

XX

PD 20-MAR-1997.

XX

XX 16-SEP-1996; 96WO-US014819.

XX

PR 15-SEP-1995; 95US-00529055.

XX

PA (UABR-) UAB RES FOUND.

XX

PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;

XX Hollingshead S, Tart R, Brooks-Walter A;

XX WPI; 1997-202002/18.

XX

PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used

XX in vaccines for protecting animals against S.pneumoniae infection.

PS Example 6; Fig 13; 296pp; English.

XX

CC This sequence shows the central portion, including the C-terminus of the
 CC alpha-helix region and some of the proline-rich region, of pneumococcal
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Bg9163.
 CC Comparison of the N-terminal and central regions (AAW14533-57 and
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
 CC be used to divide the strains into several families based on sequence
 CC homologies. PspA polypeptides, or fragments of them, can be used in
 CC vaccines to protect animals against S. pneumoniae infection and hence for
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
 CC region and the immediate 5' tip of the coding sequence are likely to be
 CC the critical sequences for predicting PspA cross-reactions and vaccine
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)
 XX
 XX

SQ Sequence 190 AA;

Query Match 97.4%; Score 453; DB 2; Length 190;
 Best Local Similarity 99.0%; Pred. No. 2e-41;
 Matches 98; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PKRIMSLSQVXLKXVCRAPIQSLDKAQAELLKLEELSGKIKELDAETAELEVLKDAE 60
 DB 10 PKRIMSLSQVXLKXVCRAPIQSLDKAQAELLKLEELSGKIKELDAETAELEVLKDAE 69
 QY 61 GNNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVDEPE 99
 DB 70 GNNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVDEPE 108

RESULT 2

ID ABW02603 standard; protein; 193 AA.

AC ABW02603;

DT 12-FEB-2004 (first entry)

DE Bg9163c pneumococcal surface protein A (PspA) central region.

KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
 KW immunological; gene therapy; immunostimulant.

OS Unidentified.

FH Key Location/Qualifiers

FT Misc-difference 1..193

FT /note= "Xaa = Unknown amino acid"

XX US6592876-B1.

XX 15-JUL-2003.

XX 15-SEP-1995; 95US-00529055.

XX 20-APR-1993; 93US-00048896.

XX 06-JUN-1995; 95US-00465746.

XX (UABR-) UAB RES FOUND.

XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;

XX WPI; 2003-862841/80.

XX Immunological composition for obtaining expression products used for
 PT detecting the presence of Streptococcus pneumoniae or its strain,
 PT comprises at least two different full length isolated gene encoding
 PT pneumococcal surface protein A.

XX Example 6; SEQ ID NO 49; 121pp; English.

XX The present invention relates to an immunological composition comprising
 CC at least 2 different full length isolated genes encoding pneumococcal
 CC surface protein A (PspAs) from different groups based on restriction

CC fragment polymorphism analysis. The invention is useful for obtaining
 CC expression products by recombinant techniques to detect, determine,
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its
 CC strain. The expression product is useful for preparing antigenic,
 CC immunological or vaccine compositions, for eliciting antibodies, an
 CC immunological response (other than or additional to antibodies) or a
 CC protective response (including antibody or other immunological response
 CC by administering compositions to a host). The invention is also useful as
 CC vaccines and in gene therapy. The present sequence is Bg9163c
 CC pneumococcal surface protein A (PspA) central region. This sequence is
 CC used in the exemplification of the invention
 XX
 XX

SQ Sequence 193 AA;

Query Match 97.4%; Score 453; DB 7; Length 193;
 Best Local Similarity 99.0%; Pred. No. 2.1e-41;
 Matches 98; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PKRIMSLSQVXLKXVCRAPIQSLDKAQAELLKLEELSGKIKELDAETAELEVLKDAE 60
 DB 13 PKRIMSLSQVXLKXVCRAPIQSLDKAQAELLKLEELSGKIKELDAETAELEVLKDAE 72
 QY 61 GNNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVDEPE 99
 DB 73 GNNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVDEPE 111

RESULT 3

ID ABU08487 standard; protein; 8991 AA.

AC ABU08487;

DT 24-JUN-2003 (first entry)

DE S. pneumoniae pneumococcal surface protein A (PspA) protein.

KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
 KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
 KW antibacterial.

OS Streptococcus pneumoniae.

FH Key Location/Qualifiers

FT Misc-difference 1..8991

FT /note= "All Xaa residues within this sequence are
 FT unknown"

XX US6500613-B1.

XX 31-DEC-2002.

XX 16-SEP-1996; 96US-00714741.

XX 15-SEP-1995; 95US-00529055.

XX (UVAL-) UNIV ALABAMA.

XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;

XX Hollingshead S, Tart R, Brooks-Walter A;

XX WPI; 2003-361534/34.

XX Isolated PspC amino acid sequence used as polymerase chain reaction or
 PT hybridization probe, comprises pneumococcal surface protein having alpha-
 PT helical, proline rich and repeat regions.

XX Disclosure; Col 145-188; 186pp; English.

XX The present invention relates to the isolation of Streptococcus
 CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
 CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-
 CC like protein having alpha-helical, proline rich and repeat regions. The

CC PspC and PspA proteins may be used in a vaccine to protect against
 CC pneumococcal infections. The polynucleotide sequences encoding PspC and
 CC PspA may be used for the expression of the proteins, and as PCR primers
 CC or hybridisation probes. The present sequence represents S. pneumoniae
 CC PspA protein
 XX
 XX Sequence 8991 AA;
 SQ
 Query Match 97.4%; Score 453; DB 6; Length 8991;
 Best Local Similarity 99.0%; Pred. No. 3e-39;
 Matches 98; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PKRMSLSQKYLKXVCRAPLQSKLDAQKAEILKLEELSGKIKELDAETAELEVLQKDAE 60
 DB 4958 PKRMSLSQKYLKXVCRAPLQSKLDAQKAEILKLEELSGKIKELDAETAELEVLQKDAE 5017
 QY 61 GNNVAYFKEGLEKTTAEKKALEXAXADLKKAVIDEPE 99
 DB 5018 GNNVAYFKEGLEKTTAEKKALEXAXADLKKAVIDEPE 5056
 RESULT 4
 AAW14574
 ID AAW14574 standard; protein; 206 AA.
 XX
 AC AAW14574;
 XX
 DT 17-OCT-2003 (revised)
 DT 28-OCT-1997 (first entry)
 XX
 XX Streptococcus pneumoniae PspA central region.
 DE
 XX
 KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
 KW bacteraemia; pneumonia.
 XX
 XX Streptococcus pneumoniae; strain Db15.
 OS
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 50 /note= "unidentified amino acid"
 FT
 FT
 PN WO9709994-AI.
 XX
 XX 20-MAR-1997.
 PD
 XX
 XX 16-SEP-1996; 96WO-US014819.
 PF
 XX
 XX 15-SEP-1995; 95US-00529055.
 PR
 XX
 PA (UABR-) UAB RES FOUND.
 XX
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
 PI Hollingshead S, Tart R, Brooks-Walter A;
 XX
 XX WPI; 1997-202002/18.
 DR
 XX
 XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
 PT in vaccines for protecting animals against S.pneumoniae infection.
 PT
 XX
 XX Example 6; Fig 13; 296pp; English.
 PS
 XX
 XX This sequence shows the central portion, including the C-terminus of the
 CC alpha-helix region and some of the proline-rich region, of pneumococcal
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Db15.
 CC Comparison of the N-terminal and central regions (AAW14533-57 and
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
 CC be used to divide the strains into several families based on sequence
 CC homologies. PspA polypeptides, or fragments of them, can be used in
 CC vaccines to protect animals against S. pneumoniae infection and hence for
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
 CC region and the immediate 5' tip of the coding sequence are likely to be
 CC the critical sequences for predicting PspA cross-reactions and vaccine

CC composition. (Updated on 17-OCT-2003 to standardise OS field)
 XX Sequence 206 AA;
 SQ
 Query Match 77.6%; Score 361; DB 2; Length 206;
 Best Local Similarity 90.2%; Pred. No. 2.8e-31;
 Matches 74; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 QY 18 RAPLQSKLDAQKAEILKLEELSGKIKELDAETAELEVLQKDAEGNNVAYFKEGLEKTT 77
 DB 18 RAPLQSKLDAQKAEILKLEELSGKIKELDAETAELEVLQKDAEGNNVAYFKEGLEKTT 77
 QY 78 AEKKALEXAXADLKKAVIDEPE 99
 DB 78 AEKKALEXAXADLKKAVIDEPE 99
 RESULT 5
 ABW02608
 ID ABW02608 standard; protein; 206 AA.
 XX
 AC ABW02608;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 DE Db15c pneumococcal surface protein A (PspA) central region.
 XX
 KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
 KW immunological; gene therapy; immunostimulant.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1..206 /note= "Xaa = Unknown amino acid"
 FT
 XX US6592876-B1.
 PN
 XX 15-JUL-2003.
 PD
 XX
 XX 15-SEP-1995; 95US-00529055.
 PF
 XX
 XX 20-APR-1993; 93US-00048896.
 PR
 XX 06-JUN-1995; 95US-00465746.
 XX
 XX (UABR-) UAB RES FOUND.
 PA
 XX
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
 PI
 XX WPI; 2003-862841/90.
 DR
 XX
 XX Immunological composition for obtaining expression products used for
 PT detecting the presence of Streptococcus pneumoniae or its strain.
 PT comprises at least two different full length isolated gene encoding
 PT pneumococcal surface protein A.
 XX
 XX Example 6; SEQ ID NO 54; 121pp; English.
 PS
 XX
 XX The present invention relates to an immunological composition comprising
 CC at least 2 different full length isolated genes encoding pneumococcal
 CC surface protein A (PspAs) from different groups based on restriction
 CC fragment polymorphism analysis. The invention is useful for obtaining
 CC expression products by recombinant techniques to detect, determine,
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its
 CC strain. The expression product is useful for preparing antigenic,
 CC immunological or vaccine compositions, for eliciting antibodies, an
 CC immunological response (other than or additional to antibodies) or a
 CC protective response (including antibody or other immunological response
 CC by administering compositions to a host). The invention is also useful as
 CC vaccines and in gene therapy. The present sequence is Db15c pneumococcal
 CC surface protein A (PspA) central region. This sequence is used in the
 CC exemplification of the invention
 XX

SQ Sequence 206 AA;
 Query Match 77.6%; Score 361; DB 7; Length 206;
 Best Local Similarity 90.2%; Pred. No. 2.8e-31;
 Matches 74; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 18 RAPIQSKLDAQKAELELLKLEELSGKIKELDAEIAELEVLQKDAEG 77
 Db KEIDESDSEDLKAEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAELEVLQKDAEG 61

QY 78 ABKAELELEXADLKKAVDEPE 99
 Db KEIDESDSEDLKAEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAELEVLQKDAEG 61

QY 78 ABKAELELEXADLKKAVDEPE 99
 Db KEIDESDSEDLKAEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAELEVLQKDAEG 61

RESULT 6
 ABW02614
 ID ABW02614 standard; protein; 170 AA.
 XX
 AC ABW02614;
 DT 12-FEB-2004 (first entry)
 DE Rct135c pneumococcal surface protein A (PspA) central region.
 XX
 KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
 KW immunological; gene therapy; immunostimulant.
 XX
 OS Unidentified.
 XX
 PN US6592876-B1.
 XX
 PD 15-JUL-2003.
 XX
 PF 15-SEP-1995; 95US-00529055.
 XX
 PR 20-APR-1993; 93US-00048896.
 PR 06-JUN-1995; 95US-00465746.
 XX
 PA (UABR-) UAB RES FOUND.
 XX
 PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
 XX WPI; 2003-862841/80.
 XX
 PT Immunological composition for obtaining expression products used for
 PT detecting the presence of Streptococcus pneumoniae or its strain,
 PT comprises at least two different full length isolated gene encoding
 PT pneumococcal surface protein A.
 XX
 PS Example 6; SEQ ID NO 60; 121pp; English.
 XX
 CC The present invention relates to an immunological composition comprising
 CC at least 2 different full length isolated genes encoding pneumococcal
 CC surface protein A (PspAs) from different groups based on restriction
 CC fragment polymorphism analysis. The invention is useful for obtaining
 CC expression products by recombinant techniques to detect, determine,
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its
 CC strain. The expression product is useful for preparing antigenic,
 CC immunological or vaccine compositions, for eliciting antibodies, an
 CC immunological response (other than or additional to antibodies) or a
 CC protective response (including antibody or other immunological response
 CC by administering compositions to a host). The invention is also useful as
 CC vaccines and in gene therapy. The present sequence is Rct135c
 CC pneumococcal surface protein A (PspA) central region. This sequence is
 CC used in the exemplification of the invention
 XX
 SQ Sequence 170 AA;
 Query Match 76.3%; Score 355; DB 7; Length 170;
 Best Local Similarity 78.6%; Pred. No. 9.8e-31;
 Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 2 KRIMSLQKVKLXXVCRAPLQSKLDAQKAELELLKLEELSGKIKELDAEIAELEVLQKDAEG 61
 Db KEIDESDSEDLKAEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAELEVLQKDAEG 61

QY 62 NNNVEAYFKEGLEKTTAEKKAEELEXADLKKAVDEPE 99
 Db NNNVEAYFKEGLEKTTAEKKAEELEXADLKKAVDEPE 99

RESULT 7
 ABW02596
 ID ABW02596 standard; protein; 181 AA.
 XX
 AC ABW02596;
 DT 12-FEB-2004 (first entry)
 DE 0922134c pneumococcal surface protein A (PspA) central region.
 XX
 KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
 KW immunological; gene therapy; immunostimulant.
 XX
 OS Unidentified.
 XX
 PN US6592876-B1.
 XX
 PD 15-JUL-2003.
 XX
 PF 15-SEP-1995; 95US-00529055.
 XX
 PR 20-APR-1993; 93US-00048896.
 PR 06-JUN-1995; 95US-00465746.
 XX
 PA (UABR-) UAB RES FOUND.
 XX
 PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
 XX WPI; 2003-862841/80.
 XX
 PT Immunological composition for obtaining expression products used for
 PT detecting the presence of Streptococcus pneumoniae or its strain,
 PT comprises at least two different full length isolated gene encoding
 PT pneumococcal surface protein A.
 XX
 PS Example 6; SEQ ID NO 42; 121pp; English.
 XX
 CC The present invention relates to an immunological composition comprising
 CC at least 2 different full length isolated genes encoding pneumococcal
 CC surface protein A (PspAs) from different groups based on restriction
 CC fragment polymorphism analysis. The invention is useful for obtaining
 CC expression products by recombinant techniques to detect, determine,
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its
 CC strain. The expression product is useful for preparing antigenic,
 CC immunological or vaccine compositions, for eliciting antibodies, an
 CC immunological response (other than or additional to antibodies) or a
 CC protective response (including antibody or other immunological response
 CC by administering compositions to a host). The invention is also useful as
 CC vaccines and in gene therapy. The present sequence is 0922134c
 CC pneumococcal surface protein A (PspA) central region. This sequence is
 CC used in the exemplification of the invention
 XX
 SQ Sequence 181 AA;
 Query Match 76.3%; Score 355; DB 7; Length 181;
 Best Local Similarity 78.8%; Pred. No. 1.1e-30;
 Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 2 KRIMSLQKVKLXXVCRAPLQSKLDAQKAELELLKLEELSGKIKELDAEIAELEVLQKDAEG 61
 Db KEIDESDSEDLKAEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAELEVLQKDAEG 61

QY 62 NNNVEAYFKEGLEKTTAEKKAEELEXADLKKAVDEPE 99
 Db NNNVEAYFKEGLEKTTAEKKAEELEXADLKKAVDEPE 99

Db 62 NNNVEAYFKEGLEKTTAEKKAELEKAEADLKKAVDPE 99

RESULT 8

ABU08489
ID ABU08489 standard; protein; 865 AA.
XX
AC ABU08489;
XX
DT 24-JUN-2003 (first entry)
XX
DE S. pneumoniae pneumococcal surface protein C (PspC) protein.
XX
KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
KW antibacterial.

OS Streptococcus pneumoniae.

XX
FH Key Location/Qualifiers
FT Peptide 1..37
FT /label= Signal_peptide
FT Protein 38..865
FT /label= Mature_PspC_protein
FT
XX

PN US6500613-B1.

XX

PD 31-DEC-2002.

XX

PF 16-SEP-1996; 96US-00714741.

XX

PR 15-SEP-1995; 95US-00529055.

XX

PA (UVAL-) UNIV ALABAMA.

XX

PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;

PI Hollingshead S, Tart R, Brooks-Walter A;

XX

DR WPI; 2003-361534/34.

XX

DR N-PSDB; ABX95377.

XX

PT Isolated PspC amino acid sequence used as polymerase chain reaction or

PT hybridization probe, comprises pneumococcal surface protein having alpha-

PT helical, proline rich and repeat regions.

XX

PS Claim 3; Fig 21; 186pp; English.

XX

CC The present invention relates to the isolation of Streptococcus

CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide

CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-

CC like protein having alpha-helical, proline rich and repeat regions. The

CC PspC and PspA proteins may be used in a vaccine to protect against

CC pneumococcal infections. The polynucleotide sequences encoding PspC and

CC PspA may be used for the expression of the proteins, and as PCR primers

CC or hybridisation probes. The present sequence represents S. pneumoniae

XX PspC protein

XX

SQ Sequence 865 AA;

XX

Query Match 76.3%; Score 355; DB 6; Length 865;

Best Local Similarity 78.6%; Pred. No. 8.1e-30;

Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 2 KRIMSLSQKVKXKVCAPLQSKLDQAOKALLKLEELSGKIKELDAEIAELEVLQKDAEG 61

Db 467 KEIDSESDYLLKEGLRAPLQSKLDTTKAKLSKLEELSDKIDELDAEIAELEVLQKDAEG 526

QY 62 NNNVEAYFKEGLEKTTAEKKAELEKAEADLKKAVDPE 99

Db 527 NNNVEAYFKEGLEKTTAEKKAELEKAEADLKKAVDPE 564

RESULT 9

AAW14593

XX AAW14593 standard; protein; 929 AA.

XX

AC AAW14593;

XX

DT 17-OCT-2003 (revised)

DT 27-OCT-1997 (first entry)

XX

DE Streptococcus pneumoniae PspC.

XX

PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;

KW bacteraemia; pneumonia.

XX

OS Streptococcus pneumoniae; strain EF6796.

XX

FH Key Location/Qualifiers

FT Peptide 1..37

FT /label= Sig_peptide

FT Protein 38..929

FT /label= Mat_protein

FT Domain 38..637

FT /label= Alpha-helix

FT Region 41..242

FT /label= Repeat_1

FT /note= "alpha-helical repeat region"

FT Region 69..637

FT /label= Coiled-coil

FT /note= "breaks in 7-residue periodicity of the coiled-

FT coil occur at amino acids 136-141, 261-304 and 383-387"

FT Region 332..492

FT /label= Repeat_2

FT /note= "alpha-helical repeat region"

FT Domain 627..689

FT /label= Proline-rich

FT Domain 913..929

FT /label= C-terminal

XX

PN W09709994-A1.

XX

PD 20-MAR-1997.

XX

PF 16-SEP-1996; 96WO-US014819.

XX

PR 15-SEP-1995; 95US-00529055.

XX

PA (UABR-) UAB RES FOUND.

XX

PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;

PI Hollingshead S, Tart R, Brooks-Walter A;

XX

DR WPI; 1997-202002/18.

XX

DR N-PSDB; AAW61728.

XX

PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used

PT in vaccines for protecting animals against S.pneumoniae infection.

XX

PS Disclosure; Fig 13; 296pp; English.

XX

CC This sequence comprises the pneumococcal protein surface C (pspC) of

CC Streptococcus pneumoniae strain EF6796. The sequence was deduced from the

CC pspC gene (AAW61728). Like PspA, PspC has an alpha-helical coiled-coil

CC region, proline-rich central region, repeat regions, with a choline

CC binding motif, and a C-terminal 17-amino acid tail. The 2 polypeptides

CC share 3 regions of high sequence identity. One is a protection-eliciting

CC region present within the alpha-helical domain. The others are the

CC proline-rich domain and a repeat domain shared with other choline-binding

CC proteins and thought to play a role in cell surface association. PspC and

CC PspA polypeptides, and their fragments, can be used in vaccines to

CC protect against S. pneumoniae infection and hence for the prevention of

CC diseases such as otitis media, meningitis, bacteraemia and pneumonia.

CC (Updated on 17-OCT-2003 to standardise OS field)

XX

SQ Sequence 929 AA;

Query Match 76.3%; Score 355; DB 2; Length 929;
 Best Local Similarity 78.6%; Pred. No. 8.9e-30;
 Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 2 KRIMSLQKVLKXVCRAPIQSKLDAQKAEILLKLEELSGIKELDAIEAEVQLKDAEG 61
 DB 531 KEIDESDSEYDKGLRAPLQSKLDTTKAKLSKLEELSDKIDELDAIEAEVQLKDAEG 590
 QY 62 NNNVEAYFKGLEKTTAEKKAELXAXADLKKAVDEPE 99
 DB 591 NNNVEAYFKGLEKTTAEKKAELXAXADLKKAVDEPE 628

RESULT 10
 AAY43384
 ID AAY43384 standard; protein; 929 AA.

XX AC AAY43384;
 XX DT 27-JAN-2000 (first entry)
 XX DE S. pneumoniae PspC protein sequence.
 XX KW PspC gene; pneumococcal surface protein C; epitope; diagnosis;
 XX KW epitopic region; immunogenic composition; vaccine composition; therapy;
 XX KW meningitis; immunological response; otitis media; bacteraemia; pneumonia;
 XX KW anti-PspA antibody.

XX OS Streptococcus pneumoniae.

XX PN W09953940-A1.

XX PD 28-OCT-1999.

XX PF 23-APR-1999; 99WO-US008895.

XX PR 23-APR-1998; 98US-0082728P.

XX PA (UYAL-) UNIV ALABAMA.

XX PI Briles DE, Hollingshead SK, Brooks-Walter A;

XX WPI; 1999-620581/53.

XX DR N-PSDB; AAZ31956.

XX PT New epitope polypeptides of Pneumococcal surface protein C, used to develop products for immunological, immunogenic or vaccine compositions, particularly against Streptococcus pneumoniae infections.

XX PS Example; Fig 11; 109pp; English.

XX CC This sequence is the Streptococcus pneumoniae pneumococcal surface protein C. The invention relates to an isolated and/or purified polypeptide comprising at least one epitope or epitopic region of Pneumococcal surface protein C (PspC). The polypeptides or vectors containing sequence encoding them can be used as immunogenic, immunological or vaccine compositions. The compositions can be used for eliciting an immunological response against Streptococcus pneumoniae (SP), which can cause otitis media, meningitis, bacteraemia and pneumonia. They can be used for eliciting an anti-PspA antibody. The nucleic acid molecules can also be used for detecting pspC, pspA or SP

XX SQ Sequence 929 AA;

Query Match 76.3%; Score 355; DB 2; Length 929;
 Best Local Similarity 78.6%; Pred. No. 8.9e-30;
 Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 2 KRIMSLQKVLKXVCRAPIQSKLDAQKAEILLKLEELSGIKELDAIEAEVQLKDAEG 61
 DB 531 KEIDESDSEYDKGLRAPLQSKLDTTKAKLSKLEELSDKIDELDAIEAEVQLKDAEG 590

QY 62 NNNVEAYFKGLEKTTAEKKAELXAXADLKKAVDEPE 99
 DB 591 NNNVEAYFKGLEKTTAEKKAELXAXADLKKAVDEPE 628

RESULT 11
 AAW14578

ID AAW14578 standard; protein; 204 AA.

XX AC AAW14578;

XX DT 17-OCT-2003 (revised)

XX DT 28-OCT-1997 (first entry)

XX DE Streptococcus pneumoniae PspA central region.

XX KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
 XX KW bacteraemia; pneumonia.

XX OS Streptococcus pneumoniae; strain Rct123.

XX FH Key Location/Qualifiers

XX FT Misc-difference 4 /note= "unidentified amino acid"

XX FT Misc-difference 8 /note= "unidentified amino acid"

XX PN W09709994-A1.

XX PD 20-MAR-1997.

XX PF 16-SEP-1996; 96WO-US014819.

XX PR 15-SEP-1995; 95US-00529055.

XX PA (UABR-) UAB RES FOUND.

XX PI Briles DE, McDaniel LS, Swiatlo E, Vother J, Crain MJ;

XX PI Hollingshead S, Tart R, Brooks-Walter A;

XX DR WPI; 1997-202002/18.

XX ST Streptococcus pneumoniae surface protein PspC and truncated PspA - used in vaccines for protecting animals against S.pneumoniae infection.

XX PS Example 6; Fig 13; 296pp; English.

XX CC This sequence shows the central portion, including the C-terminus of the alpha-helix region and some of the proline-rich region, of pneumococcal surface protein A (PspA) of Streptococcus pneumoniae strain Rct123. Comparison of the N-terminal and central regions (AAW14533-57 and AAW14562-91) of PspA polypeptides from different pneumococcal strains can be used to divide the strains into several families based on sequence homologies. PspA polypeptides, or fragments of them, can be used in vaccines to protect animals against S. pneumoniae infection and hence for the prevention of diseases such as otitis media, meningitis, bacteraemia and pneumonia. The sequence of the 3' half of the PspA alpha-helical region and the immediate 5' tip of the coding sequence are likely to be the critical sequences for predicting PspA cross-reactions and vaccine composition. (Updated on 17-OCT-2003 to standardise OS field)

XX SQ Sequence 204 AA;

Query Match 76.1%; Score 354; DB 2; Length 204;
 Best Local Similarity 86.2%; Pred. No. 1.6e-30;
 Matches 75; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 13 LKXVCRAPIQSKLDAQKAEILLKLEELSGIKELDAIEAEVQLKDAEGNNNVEAYFKEG 72
 DB 13 LKXVCRAPIQSKLDTTKAKLSKLEELSDKIDELDAIEAEVQLKDAEGNNNVEAYFKEG 72
 QY 73 LEKTTAEKKAELXAXADLKKAVDEPE 99
 |||||


```
Db 73 LEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 12
ABW02612
ID ABW02612 standard; protein; 204 AA.
XX
AC ABW02612;
XX
DT 12-FEB-2004 (first entry)
XX
DE Rct123c pneumococcal surface protein A (PspA) central region.
XX
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
KW immunological; gene therapy; immunostimulant.
XX
OS Unidentified.
XX
PH Key Location/Qualifiers
FT Misc-difference 1..204
FT /note= "Xaa = Unknown amino acid"
XX
XX US6592876-B1.
XX
PD 15-JUL-2003.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX 20-APR-1993; 93US-00048896.
XX
XX 06-JUN-1995; 95US-00465746.
XX
XX (UABR-) UAB RES FOUND.
XX
PA Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
PI WPI; 2003-862841/80.
XX
XX Immunological composition for obtaining expression products used for
PT detecting the presence of Streptococcus pneumoniae or its strain,
PT comprises at least two different full length isolated gene encoding
PT pneumococcal surface protein A.
XX
XX Example 6; SEQ ID NO 58; 121pp; English.
XX
CC The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspAs) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antigenic,
CC immunological or vaccine compositions, for eliciting antibodies, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Rct123c
CC pneumococcal surface protein A (PspA) central region. This sequence is
CC used in the exemplification of the invention
XX
SQ Sequence 204 AA;
Query Match 76.1%; Score 354; DB 7; Length 204;
Best Local Similarity 86.2%; Pred. No. 1.6e-30;
Matches 75; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 13 LKXVCRAPLQSKLDAQKAEKLLKLELSGKIKELDAIEAEVQLKDAEGNNVYAFKKG 72
DB 13 LKEGLRAPLQSKLDTTKAKLSKLELSGKIKELDAIEAEVQLKDAEGNNVYAFKKG 72
QY 73 LEKTTAEKKAELKAEADLKKAVDEPE 99
DB 73 LEKTTAEKKAELKAEADLKKAVDEPE 99

Db 73 LEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 13
AAW14580
ID AAW14580 standard; protein; 188 AA.
XX
AC AAW14580;
XX
DT 17-OCT-2003 (revised)
DT 28-OCT-1997 (first entry)
XX
DE Streptococcus pneumoniae PspA central region.
XX
KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW bacteraemia; pneumonia.
XX
OS Streptococcus pneumoniae; strain Rct135.
XX
PN WO9709994-A1.
XX
PD 20-MAR-1997.
XX
PF 16-SEP-1996; 96WO-US014819.
XX
PR 15-SEP-1995; 95US-00529055.
XX
XX (UABR-) UAB RES FOUND.
XX
PA Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX
XX WPI; 1997-202002/18.
XX
PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used
PT in vaccines for protecting animals against S.pneumoniae infection.
XX
XX Example 6; Fig 13; 296pp; English.
XX
CC This sequence shows the central portion, including the C-terminus of the
CC alpha-helix region and some of the proline-rich region, of pneumococcal
CC surface protein A (PspA) of Streptococcus pneumoniae strain Rct135.
CC Comparison of the N-terminal and central regions (AAW14533-57 and
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
CC be used to divide the strains into several families based on sequence
CC homologies. PspA polypeptides, or fragments of them, can be used in
CC vaccines to protect animals against S. pneumoniae infection and hence for
CC the prevention of diseases such as otitis media, meningitis, bacteraemia
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
CC region and the immediate 5' tip of the coding sequence are likely to be
CC the critical sequences for predicting PspA cross-reactions and vaccine
CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 188 AA;
Query Match 75.7%; Score 352; DB 2; Length 188;
Best Local Similarity 77.6%; Pred. No. 2.4e-30;
Matches 76; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

QY 2 KRMSLSQKVLKXVCRAPLQSKLDAQKAEKLLKLELSGKIKELDAIEAEVQLKDAEG 61
DB 2 KEIDSDSEDYKLEGLRAPLQSKLDTTKAKLSKLELSGKIKELDAIEAEVQLKDAEG 61
QY 62 NNNVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 99
DB 62 NNNVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPD 99

RESULT 14
ABW02613
ID ABW02613 standard; protein; 188 AA.
XX
AC ABW02613;
XX
XX 12-FEB-2004 (first entry)
XX
```

XX Rct129c pneumococcal surface protein A (PspA) central region.
XX
XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
KW immunological; gene therapy; immunostimulant.
XX
XX Unidentified.
XX
XX OS
XX PN US6592876-B1.
XX PD 15-JUL-2003.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX 20-APR-1993; 93US-00048896.
XX PR 06-JUN-1995; 95US-00465746.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX WPI; 2003-862841/80.
XX

XX Immunological composition for obtaining expression products used for
XX detecting the presence of Streptococcus pneumoniae or its strain,
XX comprises at least two different full length isolated gene encoding
XX pneumococcal surface protein A.
XX
XX Example 6; SEQ ID NO 59; 121pp; English.
XX
XX The present invention relates to an immunological composition comprising
XX at least 2 different full length isolated genes encoding pneumococcal
XX surface protein A (PspAs) from different groups based on restriction
XX fragment polymorphism analysis. The invention is useful for obtaining
XX expression products by recombinant techniques to detect, determine,
XX isolate or diagnose the presence of Streptococcus pneumoniae or its
XX strain. The expression product is useful for preparing antigenic,
XX immunological or vaccine compositions, for eliciting antibodies, an
XX immunological response (other than or additional to antibodies) or a
XX protective response (including antibody or other immunological response
XX by administering compositions to a host). The invention is also useful as
XX vaccines and in gene therapy. The present sequence is Rct129c
XX pneumococcal surface protein A (PspA) central region. This sequence is
XX used in the exemplification of the invention
XX
XX Sequence 188 AA;
SQ

Query Match 75.7%; Score 352; DB 7; Length 188;
Best Local Similarity 77.6%; Pred. No. 2.4e-30;
Matches 76; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

Qy 2 KRMSLSQKYLKXVCRAPLQSKLDQAQKAEELKLELSGKIKELDAEIAEVLQKDAEG 61
Db 2 KEIDSESDYLYKEGLRAPLQSKLDTKKAKLSLELSDKIDELDAEIAEVLQKDAEG 61
Qy 62 NNNVEAYFKEGLEKTTAEKKAELAXADLKKAVDPE 99
Db 62 NNNVEAYFKEGLEKTTAEKKAELAXADLKKAVDPE 99

RESULT 15
ABU08490
ID ABU08490 standard; protein; 1231 AA.
XX
XX AC ABU08490;
XX

XX 24-JUN-2003 (first entry)
XX

XX Fragment of S. pneumoniae PspC protein.
XX

XX Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
KW antibacterial.

XX Streptococcus pneumoniae.
XX
XX US6500613-B1.
XX
XX 31-DEC-2002.
XX
XX 16-SEP-1996; 96US-00714741.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX (UYAL-) UNIV ALABAMA.
XX
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX
XX WPI; 2003-361534/34.
XX
XX Isolated PspC amino acid sequence used as polymerase chain reaction or
XX hybridization probe, comprises pneumococcal surface protein having alpha-
XX helical, proline rich and repeat regions.
XX
XX Disclosure; Fig 22; 186pp; English.

XX The present invention relates to the isolation of Streptococcus
XX pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
XX sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-
XX like protein having alpha-helical, proline rich and repeat regions. The
XX PspC and PspA proteins may be used in a vaccine to protect against
XX pneumococcal infections. The polynucleotide sequences encoding PspC and
XX PspA may be used for the expression of the proteins, and as PCR primers
XX or hybridisation probes. The present sequence represents a fragment of S.
XX pneumoniae PspC protein

XX Sequence 1231 AA;

Query Match 74.6%; Score 347; DB 6; Length 1231;
Best Local Similarity 89.0%; Pred. No. 9.6e-29;
Matches 73; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 18 RAPLQSKLDAQKAEELKLELSGKIKELDAEIAEVLQKDAEGNNVEAYFKEGLEKTT 77
Db 509 RAPLQSKLDAQKAEELKLELSGKIKELDAEIAEVLQKDAEGNNVEAYFKEGLEKTT 568
Qy 78 AEKKAELAXADLKKAVDPE 99
Db 569 AEKKAELAXADLKKAVDPE 590

Search completed: June 21, 2005, 10:10:16
Job time : 74.8459 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:55:14 ; Search time 18.4867 Seconds
(without alignments)
399.760 Million cell updates/sec

Title: US-10-674-755-17
Perfect score: 465
Sequence: 1 PKRIMLSQKVLKXVCRAP.....KKAELXAXADLKKADEPE 99

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A COMB.pcp.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pcp.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pcp.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pcp.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pcp.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	457	98.3	99	4	US-09-147-875A-17
2	453	97.4	193	4	US-08-529-055-49
3	453	97.4	8991	4	US-08-714-741-32
4	439	94.4	99	2	US-08-710-749-16
5	373	80.2	99	4	US-09-147-875A-15
6	361	77.6	99	4	US-09-147-875A-14
7	361	77.6	206	4	US-08-529-055-54
8	359	77.2	99	2	US-08-710-749-14
9	355	76.3	170	4	US-08-529-055-60
10	355	76.3	181	4	US-08-529-055-42
11	355	76.3	864	4	US-08-714-741-40
12	354	76.1	204	4	US-08-529-055-58
13	352	75.7	99	4	US-09-147-875A-16
14	352	75.7	188	4	US-08-529-055-59
15	347	74.6	99	2	US-08-710-749-13
16	344	74.6	1231	4	US-08-714-741-41
17	344	74.0	99	2	US-08-710-749-17
18	343.5	73.9	100	4	US-09-147-875A-10
19	343.5	73.8	141	4	US-09-286-981B-2
20	343	73.8	588	4	US-08-714-741-42
21	341	73.3	99	2	US-08-710-749-15
22	341	73.1	605	4	US-08-714-741-46
23	340	65.4	99	2	US-08-710-749-11
24	304	65.4	198	4	US-08-529-055-61
25	304	65.4	619	1	US-08-465-746-2
26	304	65.4	619	1	US-08-214-164-2
27	304	65.4	619	2	US-08-467-852A-3

28	304	65.4	619	2	US-08-246-636-2	Sequence 2, Appli
29	304	65.4	619	2	US-08-247-491A-3	Sequence 2, Appli
30	304	65.4	619	2	US-08-319-795-2	Sequence 2, Appli
31	304	65.4	619	2	US-08-468-985-2	Sequence 2, Appli
32	304	65.4	619	3	US-08-312-949-2	Sequence 2, Appli
33	304	65.4	648	1	US-08-072-070-2	Sequence 2, Appli
34	304	65.4	648	1	US-08-469-434-2	Sequence 2, Appli
35	304	65.4	648	1	US-08-214-222-2	Sequence 2, Appli
36	304	65.4	648	2	US-08-467-852A-2	Sequence 2, Appli
37	304	65.4	648	2	US-08-468-718-2	Sequence 2, Appli
38	304	65.4	648	2	US-08-247-491A-2	Sequence 2, Appli
39	304	65.4	648	3	US-08-446-201-3	Sequence 3, Appli
40	304	65.4	695	1	US-08-127-499A-23	Sequence 23, Appli
41	304	65.4	695	1	US-08-482-847-23	Sequence 23, Appli
42	297	63.9	99	2	US-08-710-749-10	Sequence 10, Appli
43	297	63.9	99	4	US-09-147-875A-11	Sequence 11, Appli
44	297	63.9	204	4	US-08-529-055-51	Sequence 51, Appli
45	294	63.2	289	1	US-08-072-070-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-147-875A-17
; Sequence 17, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(99)
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid
US-147-875A-17

Query Match		98.3%	Score 457;	DB 4;	Length 99;
Best Local Similarity		100.0%	Pred. No. 2e-42;		
Matches		99;	Conservative	0;	Mismatches
				0;	Indels
				0;	Gaps
				0;	Gaps
Qy	1	PKRIMLSQKVLKXVCRAPLOSKLDAQKAE	LLKLELSGKIKELDAETAEVLQKDAE	60	
Db	1	PKRIMLSQKVLKXVCRAPLOSKLDAQKAE	LLKLELSGKIKELDAETAEVLQKDAE	60	
Qy	61	GNNVAYPKGLEKTTAEKKAELXAXADLK	KAVDEPE	99	
Db	61	GNNVAYPKGLEKTTAEKKAELXAXADLK	KAVDEPE	99	
RESULT 2					
US-08-529-055-49					
; Sequence 49, Application US/08529055					
; Patent No. 6592876					
; GENERAL INFORMATION:					
; APPLICANT: Briles, David E.					
; APPLICANT: McDaniel, Larry S.					
; APPLICANT: Swiatlo, Edwin					
; APPLICANT: Yotter, Janet					
; APPLICANT: Brooks-Walter, Alexis					
; TITLE OF INVENTION: Pneumococcal Genes, Portions					
; TITLE OF INVENTION: Thereof, Expression Products					
; TITLE OF INVENTION: Thereof, and Uses of Such Genes,					
; TITLE OF INVENTION: Portions and Products					
; NUMBER OF SEQUENCES:		73			

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 193 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-49

Query Match          97.4%; Score 453; DB 4; Length 193;
Best Local Similarity 99.0%; Pred. No. 1.2e-41;
Matches 98; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKRIMSLSQVKLVKXVCRAPIQSKLDAQKAEALLKLELSGKIKELDAEIAELEVLQKDAE 60
Db 13 PKRIMSLSQVKLVKXVCRAPIQSKLDAQKAEALLKLELSGKIKELDAEIAELEVLQKDAE 72

Qy 61 GNNVVEAYFKEGLEKTTAEKKAELXAXADLKKAVIDEPE 99
Db 73 GNNVVEAYFKEGLEKTTAEKKAELXAXADLKKAVIDEPE 111

RESULT 3
US-08-714-741-32
; Sequence 32, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a

;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8991 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-714-741-32

Query Match          97.4%; Score 453; DB 4; Length 8991;
Best Local Similarity 99.0%; Pred. No. 1.4e-39;
Matches 98; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKRIMSLSQVKLVKXVCRAPIQSKLDAQKAEALLKLELSGKIKELDAEIAELEVLQKDAE 60
Db 4958 PKRIMSLSQVKLVKXVCRAPIQSKLDAQKAEALLKLELSGKIKELDAEIAELEVLQKDAE 5017

Qy 61 GNNVVEAYFKEGLEKTTAEKKAELXAXADLKKAVIDEPE 99
Db 5018 GNNVVEAYFKEGLEKTTAEKKAELXAXADLKKAVIDEPE 5056

RESULT 4
US-08-710-749-16
; Sequence 16, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-16

Query Match          94.4%; Score 439; DB 2; Length 99;
Best Local Similarity 96.0%; Pred. No. 1.8e-40;
Matches 95; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 PKRMSLSQKVLKXVCRAPLQSKLDAQKAEILLKLELSGKIKELDAETAELEVLQKDAE 60
DB 1 PKRMSLSQKVLKXVCRAPLQSKLDAQKAEILLKLELSGKIKELDAETAELEVLQKDAE 60

QY 61 GNNVAYFKEGLEKTTAEKAELEXAXADLKKAVDEPE 99
DB 61 GNNVAYFKEGLEKTTAEKATELEXAXADLKKAVDEPE 99

RESULT 5
US-09-147-875A-15
; Sequence 15, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-15

Query Match          80.2%; Score 373; DB 4; Length 99;
Best Local Similarity 93.8%; Pred. No. 2.6e-33;
Matches 77; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 18 RAPLQSKLDAQKAEILLKLELSGKIKELDAETAELEVLQKDAEGNNVAYFKEGLEKTT 77
DB 18 RAPLQSKLDAQKAEILLKLELSGKIKELDAETAELEVLQKDAEGNNVAYFKEGLEKTT 77

QY 78 AEKKAEELEXAXADLKKAVDEPE 99
DB 78 AEKKAEELEXAEADLKKAVDEPE 99

RESULT 6
US-09-147-875A-14
; Sequence 14, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(99)
; OTHER INFORMATION: amino acid 'xaa' can be any amino acid
US-09-147-875A-14

Query Match          77.6%; Score 361; DB 4; Length 99;
Best Local Similarity 90.2%; Pred. No. 5.2e-32;
Matches 74; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 18 RAPLQSKLDAQKAEILLKLELSGKIKELDAETAELEVLQKDAEGNNVAYFKEGLEKTT 77
DB 18 RAPLQSKLDAQKAEILLKLELSGKIKELDAETAELEVLQKDAEGNNVAYFKEGLEKTT 77

QY 78 AEKKAEELEXAXADLKKAVDEPE 99
DB 78 AEKKAEELEXAEADLKKAVDEPE 99

us-10-674-755-17.ra1

QY 18 RAPLQSKLDAQKAEILLKLELSGKIKELDAETAELEVLQKDAEGNNVAYFKEGLEKTT 77
DB 18 RAPLQSKLDAQKAEILLKLELSGKIKELDAETAELEVLQKDAEGNNVAYFKEGLEKTT 77

QY 78 AEKKAEELEXAXADLKKAVDEPE 99
DB 78 AEKKAEELEXAEADLKKAVDEPE 99

RESULT 7
US-08-529-055-54
; Sequence 54, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 206 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-54

Query Match          77.6%; Score 361; DB 4; Length 206;
Best Local Similarity 90.2%; Pred. No. 1.3e-31;
Matches 74; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 18 RAPLQSKLDAQKAEILLKLELSGKIKELDAETAELEVLQKDAEGNNVAYFKEGLEKTT 77
DB 18 RAPLQSKLDAQKAEILLKLELSGKIKELDAETAELEVLQKDAEGNNVAYFKEGLEKTT 77

QY 78 AEKKAEELEXAXADLKKAVDEPE 99
DB 78 AEKKAEELEXAEADLKKAVDEPE 99

RESULT 8
US-08-710-749-14
; Sequence 14, Application US/08710749
```

; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-14

Query Match 77.2%; Score 359; DB 2; Length 99;
Best Local Similarity 90.2%; Pred. No. 8.6e-32;
Matches 74; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 18 RAPQSKLDAQKAEKLLKLEELSGKIKELDAEIAEVLQKDAEGNNVVEYFKEGLEKTT 77
Db 18 RAPQSKLDAKAKLLKLEELSGKIELEDAEIAEVLQKDVGGNNVVEYFKEGLEKTT 77

Qy 78 AEKKAELXAXADLKKAVIDEPE 99
Db 78 AEKATELEKAEADLKKAVIDEPE 99

RESULT 9
US-08-529-055-60
; Sequence 60, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY

; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-60

Query Match 76.3%; Score 355; DB 4; Length 170;
Best Local Similarity 78.8%; Pred. No. 4.6e-31;
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

Qy 2 KRIMSLSQVKVXKVCRAPIQSKLDAQKAEKLLKLEELSGKIKELDAEIAEVLQKDAEG 61
Db 2 KEIDESDSEYDLKEGLRAPLQSKLDTKAKLKKLEELSDKIDELDAEIAKLEVLQKDAEG 61

Qy 62 NNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVIDEPE 99
Db 62 NNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVIDEPE 99

RESULT 10
US-08-529-055-42
; Sequence 42, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.

```
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 181 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-42

Query Match 76.3%; Score 355; DB 4; Length 181;
Best Local Similarity 78.6%; Pred. No. 4.9e-31;
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 2 KRIMSLQKVLKXVCRAPIQSKLDAQKAELELLKLEELSGKIKELDAEIAEVLQKDAEG 61
Db 2 KEIDESSEDLKKEGLRAPLQSKLDTTKAKLSKLEELSDKIDELDAEIAEVLQKDAEG 61
QY 62 NNNVEAFKEGLEKTTAEKKAELXAXADLKKAADLKKAVDEPE 99
Db 62 NNNVEAFKEGLEKTTAEKKAELXAXADLKKAADLKKAVDEPE 99

RESULT 11
US-08-714-741-40
; Sequence 40, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 16-SEP-1996
; APPLICATION NUMBER: US/08/714,741
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Seq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 864 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 181 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-714-741-40

Query Match 76.3%; Score 355; DB 4; Length 864;
Best Local Similarity 78.6%; Pred. No. 3.4e-30;
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 2 KRIMSLQKVLKXVCRAPIQSKLDAQKAELELLKLEELSGKIKELDAEIAEVLQKDAEG 61
Db 466 KEIDESSEDLKKEGLRAPLQSKLDTTKAKLSKLEELSDKIDELDAEIAEVLQKDAEG 525
QY 62 NNNVEAFKEGLEKTTAEKKAELXAXADLKKAADLKKAVDEPE 99
Db 526 NNNVEAFKEGLEKTTAEKKAELXAXADLKKAADLKKAVDEPE 563

RESULT 12
US-08-529-055-58
; Sequence 58, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 15-SEP-1995
; APPLICATION NUMBER: US/08/529,055
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 204 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-58

Query Match 76.1%; Score 354; DB 4; Length 204;
Best Local Similarity 86.2%; Pred. No. 7.4e-31;
Matches 75; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 13 LKXVCRAPIQSKLDAQKAELELLKLEELSGKIKELDAEIAEVLQKDAEGNNNVEAFKEG 72
Db 13 LKEGLRAPLQSKLDTTKAKLSKLEELSDKIDELDAEIAEVLQKDAEGNNNVEAFKEG 72
QY 73 LEKTTAEKKAELXAXADLKKAADLKKAVDEPE 99
```

Db 73 LEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 13

US-09-147-875A-16

; Sequence 16, Application US/09147875A

; Patent No. 6638516

; GENERAL INFORMATION:

; APPLICANT: BECKER et al.

; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS

; FILE REFERENCE: 454312-2471

; CURRENT APPLICATION NUMBER: US/09/147.875A

; CURRENT FILING DATE: 1999-05-24

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 16

; LENGTH: 99

; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae

US-09-147-875A-16

Query Match 75.7%; Score 352; DB 4; Length 99;

Best Local Similarity 78.6%; Pred. No. 5e-31;

Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

Qy 2 KRIMSLQKVKLVKVCRAPLQSKLDQAQKAEELLKLEELSGKIKELDAIEAEVQLKDAEG 61

Db 2 KEIDESSEDYLYKEGERAPLQSKLDTKKAKLSKLEELSKIDELDAIEAEVQLKDAEG 61

Qy 62 NNNVEAYFKEGLEKTTAEKKAELKAELEXAXADLKKAVDEPE 99

Db 62 NNNVEAYFKEGLEKTTAEKKAELKAELEXAXADLKKAVDEPE 99

RESULT 14

US-08-529-055-59

; Sequence 59, Application US/08529055

; Patent No. 6592876

; GENERAL INFORMATION:

; APPLICANT: Briles, David E.

; APPLICANT: McDaniel, Larry S.

; APPLICANT: Swiatlo, Edwin

; APPLICANT: Yother, Janet

; APPLICANT: Brooks-Walter, Alexis

; TITLE OF INVENTION: Pneumococcal Genes, Portions

; TITLE OF INVENTION: Thereof, Expression Products

; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,

; TITLE OF INVENTION: Portions and Products

; NUMBER OF SEQUENCES: 73

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Curtis, Morris & Safford, P.C.

; STREET: 530 Fifth Avenue

; CITY: New York

; STATE: NY

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/529.055

; FILING DATE: 15-SEP-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Frommer, William S.

; REGISTRATION NUMBER: 25,506

; REFERENCE/DOCKET NUMBER: 454312-2400

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 840-3333

; TELEFAX: (212) 840-0712

; INFORMATION FOR SEQ ID NO: 59:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 188 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-529-055-59

Query Match 75.7%; Score 352; DB 4; Length 188;

Best Local Similarity 77.6%; Pred. No. 1.1e-30;

Matches 76; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

Qy 2 KRIMSLQKVKLVKVCRAPLQSKLDQAQKAEELLKLEELSGKIKELDAIEAEVQLKDAEG 61

Db 2 KEIDESSEDYLYKEGLRAPLQSKLDTKKAKLSKLEELSKIDELDAIEAEVQLKDAEG 61

Qy 62 NNNVEAYFKEGLEKTTAEKKAELKAELEXAXADLKKAVDEPE 99

Db 62 NNNVEAYFKEGLEKTTAEKKAELKAELEXAXADLKKAVDEPE 99

RESULT 15

US-08-710-749-13

; Sequence 13, Application US/08710749

; Patent No. 5955089

; GENERAL INFORMATION:

; APPLICANT: Briles, David E.

; APPLICANT: Hollingshead, Susan

; APPLICANT: Becker, Robert

; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE

; TITLE OF INVENTION: PROTEINS

; NUMBER OF SEQUENCES: 28

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Curtis, Morris & Safford

; STREET: 530 Fifth Avenue

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/710.749

; FILING DATE: 20-SEP-1996

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Frommer, William S.

; REGISTRATION NUMBER: 25,506

; REFERENCE/DOCKET NUMBER: 454312-2074

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 840-3333

; TELEFAX: (212) 840-0712

; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 99 amino acids

; TYPE: amino acid

; STRANDEDNESS: n/a

; TOPOLOGY: linear

; MOLECULE TYPE: amino acid

US-08-710-749-13

Query Match 74.6%; Score 347; DB 2; Length 99;

Best Local Similarity 86.6%; Pred. No. 1.7e-30;

Matches 71; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 18 RAPLQSKLDQAQKAEELLKLEELSGKIKELDAIEAEVQLKDAEGNNNVEAYFKEGLEKTT 77

Db 18 RAPLQSELDTKKAKLSKLEELSGKIEELDAIEAEVQLKDVGGNNNVEAYFKEGLEKTT 77

Qy 78 AEKKAELKAELEXAXADLKKAVDEPE 99

Db ||| ||| | ||||||| |||
78 AEXATELEKAEADLKXVDEPE 99

Search completed: June 21, 2005, 10:25:22
Job time : 19.4867 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 10:12:15 ; Search time 63.2388 Seconds
(without alignments)
601.118 Million cell updates/sec

Title: US-10-674-755-17

Perfect score: 465

Sequence: 1 PKRINSLQKVLKXVCRAP.....KKAELXAXADLKKADEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	457	98.3	99	15	US-10-674-755-17
2	453	97.4	193	15	US-10-299-636-64
3	373	80.2	99	15	US-10-674-755-15
4	361	77.6	99	15	US-10-674-755-14
5	361	77.6	206	15	US-10-299-636-69
6	355	76.3	170	15	US-10-299-636-75
7	355	76.3	181	15	US-10-299-636-57
8	355	76.3	643	15	US-10-299-636-95
9	355	76.3	670	9	US-09-748-875-63
10	355	76.3	670	10	US-09-298-523B-63
11	355	76.3	690	9	US-09-748-875-61

12	355	76.3	690	10	US-09-298-523B-61	Sequence 61, Appl
13	355	76.3	691	9	US-09-748-875-1	Sequence 1, Appl
14	355	76.3	691	10	US-09-298-523B-1	Sequence 1, Appl
15	355	76.3	701	9	US-09-748-875-62	Sequence 62, Appl
16	355	76.3	701	10	US-09-298-523B-62	Sequence 62, Appl
17	355	76.3	707	9	US-09-748-875-2	Sequence 2, Appl
18	355	76.3	707	10	US-09-298-523B-2	Sequence 2, Appl
19	355	76.3	711	9	US-09-748-875-3	Sequence 3, Appl
20	355	76.3	711	10	US-09-298-523B-3	Sequence 3, Appl
21	355	76.3	739	17	US-10-732-923-3294	Sequence 3294, Ap
22	355	76.3	929	9	US-09-748-875-60	Sequence 60, Appl
23	355	76.3	929	10	US-09-298-523B-60	Sequence 60, Appl
24	355	76.3	929	15	US-10-299-636-94	Sequence 94, Appl
25	354	76.1	204	15	US-10-299-636-73	Sequence 73, Appl
26	352	75.7	99	15	US-10-674-755-16	Sequence 16, Appl
27	352	75.7	188	15	US-10-299-636-74	Sequence 74, Appl
28	343.5	73.9	100	15	US-10-674-755-10	Sequence 10, Appl
29	343	73.8	141	14	US-10-254-995-2	Sequence 2, Appl
30	343	73.8	589	9	US-09-748-875-14	Sequence 14, Appl
31	343	73.8	589	10	US-09-298-523B-14	Sequence 14, Appl
32	343	73.8	589	15	US-10-299-636-97	Sequence 97, Appl
33	326.5	70.2	336	15	US-10-299-636-103	Sequence 103, App
34	304	65.4	198	15	US-10-299-636-76	Sequence 76, Appl
35	304	65.4	354	15	US-10-299-636-105	Sequence 105, App
36	304	65.4	588	15	US-10-299-636-96	Sequence 96, Appl
37	304	65.4	619	10	US-09-882-774-1	Sequence 1, Appl
38	304	65.4	619	15	US-10-282-122A-73702	Sequence 73702, A
39	304	65.4	619	16	US-10-414-532-72	Sequence 72, Appl
40	297	63.9	99	15	US-10-674-755-11	Sequence 11, Appl
41	297	63.9	204	15	US-10-299-636-66	Sequence 66, Appl
42	290.5	62.5	100	15	US-10-299-636-12	Sequence 12, Appl
43	277	59.6	195	15	US-10-299-636-86	Sequence 86, Appl
44	273	58.7	99	15	US-10-674-755-13	Sequence 13, Appl
45	215.5	46.3	100	15	US-10-674-755-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1

US-10-674-755-17
; Sequence 17, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(99)
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid
US-10-674-755-17

Query Match	98.3%	Score 457;	DB 15;	Length 99;
Best Local Similarity	100.0%;	Pred. No. 9.7e-36;		
Matches 99;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	PKRIMSLQKVLKXVCRAPLQSKLDAQKAEELKLEELSGKIKELDAETAELEVLQKDAE	60	
Db	1	PKRIMSLQKVLKXVCRAPLQSKLDAQKAEELKLEELSGKIKELDAETAELEVLQKDAE	60	
Qy	61	GNNVNEATFKEGLEKTTAEKKAELEAXADLKKADEPE	99	

D6 61 GNNVVEAYFKEGLEKTTAAEKKAELXAXADLKKAVDEPE 99

RESULT 2

US-10-299-636-64
; Sequence 64, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR FILING DATE: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae

NAME/KEY: UNSURE
LOCATION: (24)
OTHER INFORMATION: Xaa at position 24 is unknown
FEATURE:
NAME/KEY: UNSURE
LOCATION: (27)
OTHER INFORMATION: Xaa at position 27 is unknown
FEATURE:
NAME/KEY: UNSURE
LOCATION: (98)
OTHER INFORMATION: Xaa at position 98 is unknown
FEATURE:
NAME/KEY: UNSURE
LOCATION: (100)
OTHER INFORMATION: Xaa at position 100 is unknown
FEATURE:
NAME/KEY: UNSURE
LOCATION: (189)
OTHER INFORMATION: Xaa at position 189 is unknown
US-10-299-636-64

Query Match 97.4%; Score 453; DB 15; Length 193;
Best Local Similarity 99.0%; Pred. No. 5e-35;
Matches 98; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKRTMSLSQKVLKXVCRAPIQSKLDQAQKALLKLELSGKIKELDAEIAELEVLQKDAE 60

D6 13 PKRTMSLSQKVLKXVCRAPIQSKLDQAQKALLKLELSGKIKELDAEIAELEVLQKDAE 72

Qy 61 GNNVVEAYFKEGLEKTTAAEKKAELXAXADLKKAVDEPE 99

D6 73 GNNVVEAYFKEGLEKTTAAEKKAELXAXADLKKAVDEPE 111

RESULT 3

US-10-674-755-15
; Sequence 15, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471

; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-15

Query Match 80.2%; Score 373; DB 15; Length 99;
Best Local Similarity 93.9%; Pred. No. 8.5e-28;
Matches 77; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 18 RAPQSKLDQAQKALLKLELSGKIKELDAEIAELEVLQKDAEINNVEAYFKEGLEKTT 77

D6 18 RAPQSKLDQAQKALLKLELSGKIKELDAEIAELEVLQKDAEINNVEAYFKEGLEKTT 77

Qy 78 AEKKAELXAXADLKKAVDEPE 99

D6 78 AEKKAELXAXADLKKAVDEPE 99

RESULT 4

US-10-674-755-14
; Sequence 14, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)-(99)
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid
US-10-674-755-14

Query Match 77.6%; Score 361; DB 15; Length 99;
Best Local Similarity 90.2%; Pred. No. 1.2e-26;
Matches 74; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 18 RAPQSKLDQAQKALLKLELSGKIKELDAEIAELEVLQKDAEINNVEAYFKEGLEKTT 77

D6 18 RAPQSKLDQAQKALLKLELSGKIKELDAEIAELEVLQKDAEINNVEAYFKEGLEKTT 77

Qy 78 AEKKAELXAXADLKKAVDEPE 99

D6 78 AEKKAELXAXADLKKAVDEPE 99

RESULT 5

US-10-299-636-69
; Sequence 69, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan

APPLICANT: Tart, Rebecca
APPLICANT: Brooks-Walter, Alexis
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
FILE REFERENCE: 57909/361
CURRENT APPLICATION NUMBER: US/10/299,636
CURRENT FILING DATE: 2002-11-19
PRIOR APPLICATION NUMBER: 08/714,741
PRIOR FILING DATE: 1996-09-16
PRIOR APPLICATION NUMBER: 08/529,055
PRIOR FILING DATE: 1995-09-15
NUMBER OF SEQ ID NOS: 111
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 69
LENGTH: 206
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: UNSURE
LOCATION: (50)
OTHER INFORMATION: Xaa at position 50 is unknown
US-10-299-636-69

Query Match 77.6%; Score 361; DB 15; Length 206;
Best Local Similarity 90.2%; Pred. No. 2.7e-26;
Matches 74; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 18 RAPQSLDKAQAELKLESGKIKELDAEIAEVLQKDAEGNNNVEAYFKEGLEKTT 77
DB 18 RAPQSELDTKRAKLKLESGKIKELDAEIAEVLQKDAEGNNNVEAYFKEGLEKTT 77
QY 78 AEKKAELXAXADLKKAVDEPE 99
DB 78 AEKKAELXAXADLKKAVDEPE 99

RESULT 6
US-10-299-636-75
Sequence 75, Application US/10299636
Publication No. US20040077847A1
GENERAL INFORMATION:
APPLICANT: Briles, David E
APPLICANT: McDaniel, Larry S
APPLICANT: Swiatlo, Edwin
APPLICANT: Yother, Janet
APPLICANT: Crain, Marilyn J
APPLICANT: Hollingshead, Susan
APPLICANT: Tart, Rebecca
APPLICANT: Brooks-Walter, Alexis
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
FILE REFERENCE: 57909/361
CURRENT APPLICATION NUMBER: US/10/299,636
CURRENT FILING DATE: 2002-11-19
PRIOR APPLICATION NUMBER: 08/714,741
PRIOR FILING DATE: 1996-09-16
PRIOR APPLICATION NUMBER: 08/529,055
PRIOR FILING DATE: 1995-09-15
NUMBER OF SEQ ID NOS: 111
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 75
LENGTH: 170
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-10-299-636-75

Query Match 76.3%; Score 355; DB 15; Length 170;
Best Local Similarity 78.8%; Pred. No. 8e-26;
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;
QY 2 KRIMSLSQVKXKVCRAPIQSLDKAQAELKLESGKIKELDAEIAEVLQKDAEG 61
DB 2 KEIDESDSELYLKEGRAPLQSKLDTKKAELKLESGKIKELDAEIAEVLQKDAEG 61
QY 62 NNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVDEPE 99

DB 62 NNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVDEPE 99
RESULT 7
US-10-299-636-57
Sequence 57, Application US/10299636
Publication No. US20040077847A1
GENERAL INFORMATION:
APPLICANT: Briles, David E
APPLICANT: McDaniel, Larry S
APPLICANT: Swiatlo, Edwin
APPLICANT: Yother, Janet
APPLICANT: Crain, Marilyn J
APPLICANT: Hollingshead, Susan
APPLICANT: Tart, Rebecca
APPLICANT: Brooks-Walter, Alexis
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
FILE REFERENCE: 57909/361
CURRENT APPLICATION NUMBER: US/10/299,636
CURRENT FILING DATE: 2002-11-19
PRIOR APPLICATION NUMBER: 08/714,741
PRIOR FILING DATE: 1996-09-16
PRIOR APPLICATION NUMBER: 08/529,055
PRIOR FILING DATE: 1995-09-15
NUMBER OF SEQ ID NOS: 111
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 57
LENGTH: 181
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-10-299-636-57

Query Match 76.3%; Score 355; DB 15; Length 181;
Best Local Similarity 78.6%; Pred. No. 8.7e-26;
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;
QY 2 KRIMSLSQVKXKVCRAPIQSLDKAQAELKLESGKIKELDAEIAEVLQKDAEG 61
DB 2 KEIDESDSELYLKEGRAPLQSKLDTKKAELKLESGKIKELDAEIAEVLQKDAEG 61
QY 62 NNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVDEPE 99
DB 62 NNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVDEPE 99

RESULT 8
US-10-299-636-95
Sequence 95, Application US/10299636
Publication No. US20040077847A1
GENERAL INFORMATION:
APPLICANT: Briles, David E
APPLICANT: McDaniel, Larry S
APPLICANT: Swiatlo, Edwin
APPLICANT: Yother, Janet
APPLICANT: Crain, Marilyn J
APPLICANT: Hollingshead, Susan
APPLICANT: Tart, Rebecca
APPLICANT: Brooks-Walter, Alexis
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
FILE REFERENCE: 57909/361
CURRENT APPLICATION NUMBER: US/10/299,636
CURRENT FILING DATE: 2002-11-19
PRIOR APPLICATION NUMBER: 08/714,741
PRIOR FILING DATE: 1996-09-16
PRIOR APPLICATION NUMBER: 08/529,055
PRIOR FILING DATE: 1995-09-15
NUMBER OF SEQ ID NOS: 111
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 95
LENGTH: 643
TYPE: PRT
ORGANISM: Streptococcus pneumoniae

RESULT 13
US-09-748-875-1
; Sequence 1, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-1

Query Match 76.3%; Score 355; DB 9; Length 691;
Best Local Similarity 78.6%; Pred. No. 4.1e-25;
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 2 KRIMSLSQVXKLVKVCRAPIQSKLDQAQKAELELLKLEELSGKIKELDAEIAELEVLQKDAEG 61
DB 531 KEIDESDSEYDKBGLRAPLQSKLDTTKAKLSKLEELSDKIDELDAEIAELEVLQKDAEG 590

QY 62 NNNVEAYFKEGLEKTTAEKKAELKAXADLKKAADPE 99
DB 591 NNNVEAYFKEGLEKTTAEKKAELKAXADLKKAADPE 628

Search completed: June 21, 2005, 11:18:36
Job time : 63.2388 secs

RESULT 14
US-09-298-523B-1
; Sequence 1, Application US/09298523B
; Publication No. US20030059438A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/298,523B
; CURRENT FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-1

Query Match 76.3%; Score 355; DB 10; Length 691;
Best Local Similarity 78.6%; Pred. No. 4.1e-25;
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 2 KRIMSLSQVXKLVKVCRAPIQSKLDQAQKAELELLKLEELSGKIKELDAEIAELEVLQKDAEG 61
DB 531 KEIDESDSEYDKBGLRAPLQSKLDTTKAKLSKLEELSDKIDELDAEIAELEVLQKDAEG 590

QY 62 NNNVEAYFKEGLEKTTAEKKAELKAXADLKKAADPE 99
DB 591 NNNVEAYFKEGLEKTTAEKKAELKAXADLKKAADPE 628

RESULT 15
US-09-748-875-62
; Sequence 62, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:53:39 ; Search time 9.9 Seconds
(without alignments)
962.168 Million cell updates/sec

Title: US-10-674-755-17

Perfect score: 465

Sequence: 1 PIRMSLSQVXKXVCRAP.....KKAEXAXADLKKAQVDEPE 99

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	304	65.4	619	2 A97887	surface protein ps
2	304	65.4	619	2 A41971	surface protein ps
3	110.5	23.8	744	2 P95013	pneumococcal surfa
4	103	22.2	161	2 S48396	tropomyosin TM2 -
5	103	22.2	281	2 F75216	hypothetical prote
6	101.5	21.8	1006	2 C70445	ATPase subunit of
7	100	21.5	1312	2 T30845	probable DNA repai
8	95	20.4	785	2 T01025	hypothetical prote
9	94	20.2	229	2 S70532	outer surface prot
10	94	20.2	1808	2 T15099	hypothetical prote
11	93	20.0	876	2 A23767	myosin heavy chain
12	92	19.8	395	2 AC1754	capsid protein [ba
13	92	19.8	1171	2 T45706	chromosome-associ
14	91.5	19.7	852	2 D72230	conserved hypotet
15	91.5	19.7	1269	2 P84730	probable myosin he
16	91	19.6	880	2 F75103	conserved hypotet
17	90.5	19.5	1976	2 A59252	myosin heavy chain
18	90	19.4	1319	2 A28313	glued protein - fr
19	90	19.4	1938	2 A59293	skeletal myosin he
20	89.5	19.2	399	2 E71169	hypothetical prote
21	89	19.1	784	2 T05409	hypothetical prote
22	88.5	19.0	2116	2 A26655	myosin heavy chain
23	88.5	19.0	3187	2 JC5837	364K Golgi complex
24	88	18.9	955	2 S24348	myosin heavy chain
25	88	18.9	1938	1 JX0178	myosin heavy chain
26	88	18.9	1940	2 A29320	myosin heavy chain
27	88	18.9	1974	2 T30010	hypothetical prote
28	88	18.9	4574	2 G02520	plectin - human
29	87.5	18.8	279	2 D71453	hypothetical prote

RESULT 1

A97887

surface protein pspA precursor [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: A97887
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczyk, L.; Burgett, S.; DeHoff, B.S.; E.
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: A97887
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-619 <KUR>
A:Cross-References: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:AE007317; PIDN:AAK98925.1; PID:gl
C:Genetics:
A:Gene: pspA

Query Match 65.4%; Score 304; DB 2; Length 619;

Best Local Similarity 80.5%; Pred. No. 3.8e-16;

Matches 66; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 18 RAPLOSKLDAQKAEILLKLEELSGKIKELDAEIAELVQLKDAEGNNVAYPKGLEKTT 77

Db 240 RAPLOSKLDAKAKLKLKLEELSDKIDELDAEIAKLEQDKAAAEENNVEDYFKGLEKTI 299

QY 78 AEKKAEXAXADLKKAQVDEPE 99

Db 300 AAKKAELEKTEADLKKAQVNEPE 321

RESULT 2

A41971

surface protein pspA precursor - Streptococcus pneumoniae
N:Alternate names: pneumococcal surface protein A
C:Species: Streptococcus pneumoniae
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A41971; A60282; A33134
R:Yother, J.; Briles, D.E.
J. Bacteriol. 174, 601-609, 1992

A:Title: Structural properties and evolutionary relationships of PspA, a surface protein
A:Reference number: A41971; MUID:92105030; PMID:1729249
A:Accession: A41971
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-619 <VOT>

A:Cross-References: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:M74122; NID:G153840; PIDN:AAA2701

A>Note: sequence extracted from NCBI backbone (NCBIN:75635, NCBIIP:75636)

R:Talkington, D.F.; Crimmins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.

C;Accession: T01025
R;Theologas, A.; Vysotskaia, V.S.; Osborne, B.I.; Schwartz, J.R.; Federspiel, N.A.; Kwan,
Oefner, P.; Davis, R.W.
submitted to the EMBL Data Library, May 1998
A;Description: Arabidopsis thaliana chromosome 1 YAC YUP8H12R sequence.
A;Reference number: Z14227
A;Accession: T01025
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-785 <THE>
A;Cross-references: UNIPROT:O64521, EMBL:AC002986, NID:g2494106, PID:g3152577, GSPDB:GNOC
A;Experimental source: cultivar Columbia
C;Genetics:
A;Gene: ATSP:YUP8H12R.8
A;Map position: 1
A;Introns: 108/3; 157/3; 190/3; 341/3; 394/3; 420/3; 455/2; 495/3; 526/3; 543/1; 577/3; 5

Query Match 20.4%; Score 95; DB 2; Length 785;
Best Local Similarity 26.0%; Pred. No. 5;
Matches 32; Conservative 25; Mismatches 36; Indels 30; Gaps 4;

Qy 2 KRIMSLSQVKLKXVCRAPLOSQKLDQAQAEILLKLE-----ELSGRIKEL 45
Db 59 QKYLSSQDSSLESQNAKLSDFDRLAELAQAQAKHQHLQSVRPCVIEKDGEVERM 118
Qy 46 DAETAELE-----VOLKDA---EGNNVNVEAYFKE--GLEKTTAAKKAELEXAXADL 91
Db 119 STEMSELHKSRQLMWELLEQDAISEKNSTIKSYLDKIVKLTDTTSSEKEARLAETAEL 178
Qy 92 KKA 94
Db 179 ARS 181

RESULT 9
S70532
outer surface protein F precursor - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S70532
R;Akins, D.R.; Porcella, S.F.; Popova, T.G.; Shevchenko, D.; Baker, S.I.; Li, M.; Norgard
Mol. Microbiol. 18, 507-520, 1995
A;Title: Evidence for in vivo but not in vitro expression of a Borrelia burgdorferi outer
A;Reference number: S70531; MUID:96342380; PMID:8748034
A;Accession: S70532
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-229 <AKI>
A;Cross-references: UNIPROT:Q44735; EMBL:U19754; NID:g3318660; PIDN:AAC26147.1; PID:g896(
C;Genetics:
A;Gene: ospF
C;Superfamily: outer surface protein F ospF
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-229/Product: outer surface protein F #status predicted <MAT>

Query Match 20.2%; Score 94; DB 2; Length 229;
Best Local Similarity 29.3%; Pred. No. 1.8;
Matches 29; Conservative 15; Mismatches 27; Indels 28; Gaps 3;

Qy 2 KRIMSLSQVKLKXVCRAPLOSQKLDQAQAEILLKLEELSGRIKELDAEI---AELEVQLKD 58
Db 76 KRIQLAKEKI-----EKLEAKTSLTKTYSEYEKLAQIKEKLGKADLEDKJK- 123
Qy 59 AEGNNVNVEAYFKEGLEKTTAAKKAELEXAXADLKKAVDE 97
Db 124 -----GLDSLSKKKKERKKALEDADAKKKFEE 149

RESULT 10
T15099
hypothetical protein W03F8.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text change 09-Jul-2004

R.Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.

Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing
A:Reference number: A72200; PMID:99287316; PMID:10360571

A:Accession: D72230
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-852 <ARN>
A:Cross-references: UNIPROT:Q9X1X1; GB:AE001806; GB:AE000512; NID:g4982196; PIDN:AAD3670
A:Experimental source: strain MSB8

C:Genetics:
A:Gene: Tm1636

C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 19.7%; Score 91.5; DB 2; Length 852;
Best Local Similarity 31.1%; Pred. No. 10;
Matches 23; Conservative 18; Mismatches 22; Indels 11; Gaps 2;

QY 21 LOSKLDQAQKAELEKLE---LSGKIKELDAEIAEVLQKDAEGNNVVEAYFKEGLEKT 76

DB 521 LEEKLDEKRLKRIEERHSISQKITAADVQISQIENQLKEIKGE-----IEAKRET 573

QY 77 TAEKKAELXAXAD 90

DB 574 LKEQREMDQLKSD 587

RESULT 15

F84730

probable myosin heavy chain [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: F84730

R.Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; PMID:20083487; PMID:10617197

A:Accession: F84730

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1269 <STO>

A:Cross-references: GB:AE002093; NID:g6598483; PIDN:AAC69932.2; GSPDB:GN00139

C:Genetics:

A:Gene: At2g32240

A:Map position: 2

Query Match 19.7%; Score 91.5; DB 2; Length 1269;
Best Local Similarity 27.0%; Pred. No. 15;
Matches 27; Conservative 22; Mismatches 48; Indels 3; Gaps 1;

QY 3 RIMSLQKVKLVKVCRAPIQSKLDQAQKAELEKLEELSGKIKELDAEIAEVLQKDAEGN 62

DB 300 RLLETRQKVSSTEALIDELTQELQKASESRPFKEELSVLQDLDAQTGKLQAKLSQEIGI 359

QY 63 NN---VEAYFKEGLEKTAETKKAELXAXADLKAVDEPE 99

DB 360 NSKLAEELKEELLESLSKQDEKRLTANEKLAELVLEKE 399

Search completed: June 21, 2005, 10:12:03
Job time : 10.9 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:54:24 ; Search time: 61.3194 Seconds
(without alignments)
826.751 Million cell updates/sec

Title: US-10-674-755-17
Perfect score: 465
Sequence: 1 PRRMSLSQVKKXVCRAP.....KKAELEXAXADLKKAUDEPE 99

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	373	80.2	395	2	Q9LAY2	Q9lay2 streptococc
2	373	80.2	408	2	Q9LAY0	Q9lay0 streptococc
3	367	78.9	249	2	Q9L575	Q9l575 streptococc
4	366	78.7	426	2	Q9LAY5	Q9lay5 streptococc
5	365	78.5	224	2	Q8GNS8	Q8gns8 streptococc
6	362	77.8	99	2	Q8KQK4	Q8kqk4 streptococc
7	355	76.3	739	2	Q9RQT4	Q9rqt4 streptococc
8	355	76.3	820	2	Q9RQT1	Q9rqt1 streptococc
9	355	76.3	929	2	Q9KK19	Q9kk19 streptococc
10	355	76.3	929	2	Q9ZAY5	Q9zay5 streptococc
11	336	72.3	437	2	Q9LAY4	Q9lay4 streptococc
12	315.5	67.8	869	2	Q9KK27	Q9kk27 streptococc
13	304	65.4	619	2	Q54972	Q54972 streptococc
14	304	65.4	619	2	Q8DR10	Q8dr10 streptococc
15	297	63.9	417	2	Q9LAY3	Q9lay3 streptococc
16	282	60.6	415	2	Q9LAY1	Q9lay1 streptococc
17	221.5	47.6	246	2	Q9L578	Q9l578 streptococc
18	219.5	47.2	194	2	Q9L5B5	Q9l5b5 streptococc
19	219.5	47.2	218	2	Q6UEB2	Q6ueb2 streptococc
20	219.5	47.2	233	2	Q9L568	Q9l568 streptococc
21	219.5	47.2	236	2	Q9L569	Q9l569 streptococc
22	219.5	47.2	243	2	Q9L564	Q9l564 streptococc
23	219.5	47.2	243	2	Q9L567	Q9l567 streptococc
24	219.5	47.2	244	2	Q9L565	Q9l565 streptococc
25	219.5	47.2	249	2	Q9L566	Q9l566 streptococc
26	219.5	47.2	249	2	Q9L570	Q9l570 streptococc
27	219.5	47.2	254	2	Q9L563	Q9l563 streptococc
28	219.5	47.2	401	2	Q9LAZ2	Q9laz2 streptococc
29	217.5	46.8	222	2	Q9L577	Q9l577 streptococc
30	217.5	46.8	262	2	Q9L576	Q9l576 streptococc
31	217.5	46.8	415	2	Q9LAY7	Q9lay7 streptococc

32	216.5	46.6	225	2	Q9L591	Q9l591 streptococc
33	215.5	46.3	394	2	Q9LAY6	Q9lay6 streptococc
34	215.5	46.3	395	2	Q9LAZ1	Q9laz1 streptococc
35	214.5	46.1	393	2	Q9LAZ3	Q9laz3 streptococc
36	210.5	45.3	416	2	Q9LAY8	Q9lay8 streptococc
37	209.5	45.1	255	2	Q9L581	Q9l581 streptococc
38	209.5	45.1	255	2	Q9L5B6	Q9l5b6 streptococc
39	206.5	44.4	406	2	Q9LAZ0	Q9laz0 streptococc
40	203.5	43.8	237	2	Q9L592	Q9l592 streptococc
41	203.5	43.8	395	2	Q9LAY9	Q9lay9 streptococc
42	198.5	42.7	340	2	Q8KQK5	Q8kqk5 streptococc
43	194.5	41.8	207	2	Q8GNS9	Q8gns9 streptococc
44	160	34.4	107	2	Q8KQK2	Q8kqk2 streptococc
45	160	34.4	246	2	Q9L5B4	Q9l5b4 streptococc

ALIGNMENTS

RESULT 1
Q9LAY2
ID Q9LAY2 PRELIMINARY; PRT; 395 AA.
AC Q9LAY2;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE PspA (Fragment).
OS Name=pspA;
GN Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF6796;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071813; AAF27709.1; -;
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin_abc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 395
SQ SEQUENCE 395 AA; 42963 MW; 58B6EF956BCBCC1E CRC64;

Query Match 80.2%; Score 373; DB 2; Length 395;
Best Local Similarity 93.9%; Pred. No. 1.4e-19;
Matches 77; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 18 RAPLSKLDQAQKAEILKLELSGKIKELDAEIALEVLQKDAEGNNNNVAYFKEGLEKTT 77
DB 242 RAPLSKLDQAQKAEILKLELSGKIKELDAEIALEVLQKDAEGNNNNVAYFKEGLEKTT 301
QY 78 AEKKAEXAXADLKKAUDEPE 99
DB 302 AEKKAEXAXADLKKAUDEPE 323

RESULT 2
Q9LAY0
ID Q9LAY0 PRELIMINARY; PRT; 408 AA.
AC Q9LAY0;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

```
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9163;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071815; AAF27711.1; -.
DR InterPro; IPR009053; Pfefoldin.
DR InterPro; IPR011047; Quin_alc_DH_like.
FT NON_TER 408 408
FT SEQUENCE 408 AA; 44254 MW; 4F64D874217297EF CRC64;

Query Match 80.2%; Score 373; DB 2; Length 408;
Best Local Similarity 93.9%; Pred. No. 1.4e-19;
Matches 77; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 18 RAPQSLDQAQKALLKLEELSGKIKELDAIEAEVQLKDAEGNNVVEAYFKGLEKTT 77
Db 245 RAPQSLDQAQKALLKLEELSGKIEELDAIEAEVQLKDAEGNNVVEAYFKGLEKTT 304

Qy 78 AEKKAELXAXADLKKAVDPE 99
Db 305 AEKKAELKAEADLKKAVDPE 326

RESULT 3
Q9L575 PRELIMINARY; PRT; 249 AA.
AC Q9L575;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=195;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
pneumococcal strains in the United States and of internationally
disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=195;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255552; AAF68105.1; -.
FT NON_TER 249 249
FT NON_TER 249 249
FT SEQUENCE 249 AA; 26986 MW; 7916D5014E387BD8 CRC64;

Query Match 78.9%; Score 367; DB 2; Length 249;
Best Local Similarity 79.6%; Pred. No. 2.4e-19;
Matches 78; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

Qy 2 KRIMSLSQKVLKXVCRAPLQSLDKAQAELKLEELSGKIKELDAIEAEVQLKDAEG 61
Db 75 KEIDSDSDYIKGFRAPLQSLDTKKALKLEELSGKIEELDAIEAEVQLKDAEG 134

Qy 62 NNVVEAYFKGLEKTTAEKKAELXAXADLKKAVDPE 99
Db 135 NNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDPE 172

us-10-674-755-17.rup

RESULT 4
Q9LAY5 PRELIMINARY; PRT; 426 AA.
AC Q9LAY5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBL5;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071810; AAF27706.1; -.
DR HSP; P00192; IM6T.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 426 426
FT SEQUENCE 426 AA; 46534 MW; 81AA11348CBE6634 CRC64;

Query Match 78.7%; Score 366; DB 2; Length 426;
Best Local Similarity 87.4%; Pred. No. 4.8e-19;
Matches 76; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 13 LKXVCRAPLQSLDKAQAELKLEELSGKIKELDAIEAEVQLKDAEGNNVVEAYFKEG 72
Db 227 VKESLRAPLQSLDTKKALKLEELSGKIEELDAIEAEVQLKDAEGNNVVEAYFKEG 286

Qy 73 LEKTTAEKKAELXAXADLKKAVDPE 99
Db 287 LEKTTAEKKAELKAEADLKKAVDPE 313

RESULT 5
Q8GNS8 PRELIMINARY; PRT; 224 AA.
AC Q8GNS8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PN124;
RX MEDLINE=2241996; PubMed=12354862;
RA Dicuonzo G., Gherardi G., Gertz R.E., D'Ambrosio F., Goglio A.,
RA Lorino G., Recchia S., Pantosti A., Beall B.;
RT "genotypes of invasive pneumococcal isolates recently recovered from
Italian patients.";
RL J. Clin. Microbiol. 40:3660-3665(2002).
DR EMBL; AF490267; AAN37735.1; -.
DR HSP; P00192; IAPC.
DR InterPro; IPR009082; His_kin_homodim.
FT NON_TER 224 224
FT NON_TER 224 224
FT SEQUENCE 224 AA; 23418 MW; 48674E27AFB66A95 CRC64;
```



```
Query Match      78.5%; Score 365; DB 2; Length 224;
Best Local Similarity 91.5%; Pred. No. 3.le-19;
Matches 75; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 18 RAPQSKLDAQKAEELKLELSKIKELDAEAEVQLKDAEGNNVAYFKEGLEKTT 77
DB 34 RAPQSELDTKKAKLKLSELSKIKELDAEAEVQLKDAEGNNVAYFKEGLEKTT 93
QY 78 AEKAELEXAXADLKXAVDEPE 99
DB 94 AEKAELEKAEADLKXAVDEPE 115

RESULT 6
Q8KQK4 PRELIMINARY; PRT; 99 AA.
AC Q8KQK4;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DE Pneumococcal surface protein A (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=371/00;
RX MEDLINE=22170754; PubMed=12183557;
RX DOI=10.1128/IAI.70.9.5086-5090.2002;
RA Miyaji E.N., Ferreira D.M., Lopes A.P.Y., Brandileone M.C.C.,
RA Dias W.O., Leite L.C.C.;
RT "Analysis of serum cross-reactivity and cross-protection elicited by
RT immunization with DNA vaccines against Streptococcus pneumoniae
RT expressing PspA fragments from different clades.";
RL Infect. Immun. 70:5086-5090(2002).
DR EMBL; AY082388; AAL92493.1; -.
FT NON_TER 1
FT NON_TER 99
SQ SEQUENCE 99 AA; 11105 MW; 7A13308CA174A3A7 CRC64;

Query Match      77.8%; Score 362; DB 2; Length 99;
Best Local Similarity 78.6%; Pred. No. 2.3e-19;
Matches 77; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 2 KRIMSLSQVKLVKVCRAPLQSKLDAQKAEELKLELSKIKELDAEAEVQLKDAEG 61
DB 2 KEIDESDSDYKLGRLAPLQSELDTKKAKLLKLELSKIKELDAEAEVQLKDAEG 61
QY 62 NNNVEAYFKEGLEKTTAEKKAELXAXADLKXAVDEPE 99
DB 62 NNNVEAYFKEGLEKTTAEKKAELXAXADLKXAVDEPE 99

RESULT 7
Q9ROT4 PRELIMINARY; PRT; 739 AA.
AC Q9ROT4;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Hypothetical protein pspC (Fragment).
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL34;
RX MEDLINE=20038319; PubMed=10569772;

Query Match      76.3%; Score 355; DB 2; Length 820;
Best Local Similarity 78.6%; Pred. No. 5.6e-18;
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;
```

```
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, PspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; AF068647; AAF13457.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW_binding_1; 1.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
KW Hypothetical protein.
FT NON_TER 739
SQ SEQUENCE 739 AA; 83960 MW; 7EB2F2F676ABF989 CRC64;

Query Match      76.3%; Score 355; DB 2; Length 739;
Best Local Similarity 78.6%; Pred. No. 5.1e-18;
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 2 KRIMSLSQVKLVKVCRAPLQSKLDAQKAEELKLELSKIKELDAEAEVQLKDAEG 61
DB 538 KEIDESDSDYKLGRLAPLQSKLDTKKAELKLELSKIKELDAEAEVQLKDAEG 597
QY 62 NNNVEAYFKEGLEKTTAEKKAELXAXADLKXAVDEPE 99
DB 598 NNNVEAYFKEGLEKTTAEKKAELXAXADLKXAVDEPE 635

RESULT 8
Q9ROT1 PRELIMINARY; PRT; 820 AA.
AC Q9ROT1;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Hypothetical protein pspC (Fragment).
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9163;
RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, PspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; AF068650; AAF13460.1; -.
DR HSSP; P04268; IIC2.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW_binding_1; 1.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
KW Hypothetical protein.
FT NON_TER 820
SQ SEQUENCE 820 AA; 91752 MW; 33C095849ABB0942 CRC64;

Query Match      76.3%; Score 355; DB 2; Length 820;
Best Local Similarity 78.6%; Pred. No. 5.6e-18;
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;
```

```
Qy 2 KRMSLSQKVKLKKVCRAPIQSKLDAQKAEALLKLEELSGKIKELDAEIAEVLQKDAEG 61
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 531 KEIDESDSYDLKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAEVLQKDAEG 590

Qy 62 NNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVDEPE 99
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 591 NNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVDEPE 628

RESULT 9
Q9KK19 ID Q9KK19 PRELIMINARY; PRT; 929 AA.
AC Q9KK19
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein pspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8710;
RX MEDLINE=2188621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
Streptococcus pneumoniae.";
RL Gene 284:63-71(2002).
DR EMBL; AF154037; AAF73809.1; -.
DR HSP; P06653; 1H8G.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR009053; Prefoldin.
DR Pfam; PF01473; CW_binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; YsIRK_signal; 1.
DR TIGRFAMs; TIGR01168; YsIRK_signal; 1.
SQ SEQUENCE 929 AA; 105003 MW; 2DC8293302FAFA64 CRC64;

Query Match 76.3%; Score 355; DB 2; Length 929;
Best Local Similarity 78.6%; Pred. No. 6.4e-18;
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

Qy 2 KRMSLSQKVKLKKVCRAPIQSKLDAQKAEALLKLEELSGKIKELDAEIAEVLQKDAEG 61
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 531 KEIDESDSYDLKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAEVLQKDAEG 590

Qy 62 NNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVDEPE 99
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 591 NNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVDEPE 628

RESULT 10
Q9ZAY5 ID Q9ZAY5 PRELIMINARY; PRT; 929 AA.
AC Q9ZAY5;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein C.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bf6796;
RX MEDLINE=20038319; PubMed=10569772;
```

```
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
protein, PspC, which elicits cross-reactive antibodies to PspA and
provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; U72655; AAD00184.1; -.
DR HSP; P06653; 1HCX.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW_binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; YsIRK_signal; 1.
DR TIGRFAMs; TIGR01168; YsIRK_signal; 1.
SQ SEQUENCE 929 AA; 104991 MW; 2DC8293302FFB081 CRC64;

Query Match 76.3%; Score 355; DB 2; Length 929;
Best Local Similarity 78.6%; Pred. No. 6.4e-18;
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

Qy 2 KRMSLSQKVKLKKVCRAPIQSKLDAQKAEALLKLEELSGKIKELDAEIAEVLQKDAEG 61
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 531 KEIDESDSYDLKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAEVLQKDAEG 590

Qy 62 NNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVDEPE 99
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 591 NNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVDEPE 628

RESULT 11
Q9LAY4 ID Q9LAY4 PRELIMINARY; PRT; 437 AA.
AC Q9LAY4
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E134;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071811; AAF27707.1; -.
DR InterPro; IPR002479; CW_binding.
DR Pfam; PF01473; CW_binding_1; 1.
DR NON TER 437
SQ SEQUENCE 437 AA; 48071 MW; F6EFD2CD13E08CD8 CRC64;

Query Match 72.3%; Score 336; DB 2; Length 437;
Best Local Similarity 86.6%; Pred. No. 7.6e-17;
Matches 71; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 18 RAPLQSKLDAQKAEALLKLEELSGKIKELDAEIAEVLQKDAEGNNNVEAYFKEGLEKTT 77
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 252 RAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAEVLQKDAEGNNNVEAYFKEGLEKTT 311

Qy 78 AEKKAELXAXADLKKAVDEPE 99
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 312 AEKKAELXAXADLKKAVDEPE 333

RESULT 12
```

```
Q9KK27
ID Q9KK27 PRELIMINARY; PRT; 869 AA.
AC Q9KK27;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Surface protein PspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]_SEQUENCE FROM N.A.
RP STRAIN=95;
RC MEDLINE=21898621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae.";
RL Gene 284:63-71(2002).
DR EMBL; AF154032; AAF73801.1; -.
DR HSP; P06653; 1H8G.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW binding 1; 8.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRfam; TIGR01168; Ysirk signal; 1.
SQ SEQUENCE 869 AA; 98732 MW; AFF2B504347E0220 CRC64;

Query Match 67.8%; Score 315.5; DB 2; Length 869;
Best Local Similarity 72.4%; Pred. No. 4.6e-15;
Matches 71; Conservative 3; Mismatches 23; Indels 1; Gaps 1;

QY 2 KRIMLSQKXKVCRAPOSKLDAQKAELELLESGKIKELDAEIAEVLQKDAEG 61
DB 538 KEIDESDSYLLKEGURAPQSKLDTPKKAKLSKLELSDKIDELVNCNLSQKDAEG 596

QY 62 NNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVIDEPE 99
DB 597 NNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVIDEPE 634

RESULT 13
Q54972
ID Q54972 PRELIMINARY; PRT; 619 AA.
AC Q54972;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pneumococcal surface protein A precursor.
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]_SEQUENCE FROM N.A.
RP MEDLINE=92105030; PubMed=1729249;
RX Yother J., Briles D.E.;
RT "Structural properties and evolutionary relationships of PspA, a
RT surface protein of Streptococcus pneumoniae, as revealed by sequence
RT analysis.";
RL J. Bacteriol. 174:601-609(1992).
RN [2]_SEQUENCE FROM N.A.
RP Yother J., Briles D.E.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M74122; AAA27018.1; -.
DR PIR; A41971; A41971.
DR PIR; A97887; A97887.
```

```
DR HSP; P06653; 1HCX.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam; PF01473; CW binding 1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_3.
KW SIGNAL.
FT SIGNAL. 1 31 Potential.
FT CHAIN 32 619 pneumococcal surface protein A.
SQ SEQUENCE 619 AA; 68605 MW; 5AA8BDB40C2841CA CRC64;

Query Match 65.4%; Score 304; DB 2; Length 619;
Best Local Similarity 80.5%; Pred. No. 2.3e-14;
Matches 66; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 18 RAPLOSKLDAQKAELELLESGKIKELDAEIAEVLQKDAEGNNVYAYFKEGLEKTT 77
DB 240 RAPLOSKLDAQKAELELLESGKIKELDAEIAEVLQKDAEGNNVYAYFKEGLEKTT 299

QY 78 AEKKAELXAXADLKKAVIDEPE 99
DB 300 AAKKAELXAXADLKKAVIDEPE 321

RESULT 14
Q8DR10
ID Q8DR10 PRELIMINARY; PRT; 619 AA.
AC Q8DR10;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein pspA.
GN Name=pspA; OrderedLocNames=sp0121;
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=171101;
RN [1]_TaxID=171101;
RP SEQUENCE FROM N.A.
RX MEDLINE=21429245; PubMed=11544234;
RX DOI=10.1128/JB.183.19.5709-5717.2001;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Barget S.,
RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAhren S.N., McHenry M., McLeaster K., Mundy C.W., Niclas T.I.,
RA Norris P.H., O'Gara M., Peery R.B., Robertson G.T., Rokey P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,
RA Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717(2001).
DR EMBL; AE008396; AAK98925.1; -.
DR PIR; A41971; A41971.
DR PIR; A97887; A97887.
DR HSP; P06653; 1HCX.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam; PF01473; CW binding 1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_3.
KW Complete proteome.
SQ SEQUENCE 619 AA; 68605 MW; 5AA8BDB40C2841CA CRC64;

Query Match 65.4%; Score 304; DB 2; Length 619;
Best Local Similarity 80.5%; Pred. No. 2.3e-14;
Matches 66; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 18 RAPLOSKLDAQKAELELLESGKIKELDAEIAEVLQKDAEGNNVYAYFKEGLEKTT 77
DB 240 RAPLOSKLDAQKAELELLESGKIKELDAEIAEVLQKDAEGNNVYAYFKEGLEKTT 299

QY 78 AEKKAELXAXADLKKAVIDEPE 99
DB 300 AAKKAELXAXADLKKAVIDEPE 321
```

```
RESULT 15
Q9LAY3 PRELIMINARY; PRT; 417 AA.
AC Q9LAY3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF10197;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K.; Becker R.; Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071812; AAF27708.1; -.
DR HSP; P00192; 256B.
FT NON TER 417
SQ SEQUENCE 417 AA; 46960 MW; 876EAD3276506EEC CRC64;

Query Match 63.9%; Score 297; DB 2; Length 417;
Best Local Similarity 68.4%; Pred. No. 5.le-14;
Matches 67; Conservative 8; Mismatches 23; Indels 0; Gaps 0;

Qy 2 KRIMSLSQVXLKXVCAPLQSLDQAELLKLEELSGKIKELDAEIAELEVLQKDAEG 61
Db 214 KEIDSESDYVXGEGFAPLQSELDQAQKLSLEELSDKIDELDAEIAKLELDQKAAEE 273

Qy 62 NNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVIDPE 99
Db 274 NNNVEDYFKEGLEKTIKKALEKTEADLKKAVIDPE 311
```

Search completed: June 21, 2005, 10:22:13
Job time : 62.3194 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:50:39 ; Search time 76.0837 Seconds
(without alignments)
518.502 Million cell updates/sec

Title: US-10-674-755-18

Perfect score: 505

Sequence: 1 LAKQTELEKLLDLDEPKT.....TQKELDALNELGPDGDEE 102

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	469	92.9	213	7	ABW02601 Bg8090c p
2	469	92.9	8991	6	ABU08487 S. pneumo
3	468	92.7	416	8	ADK52498 alpha hel
4	468	92.7	526	8	ADK52497 PspA mole
5	468	92.7	744	6	ABU00449 S. pneumo
6	468	92.7	744	8	ADM52054 S. pneumo
7	468	92.7	745	3	AAY81652 Streptoco
8	467	92.5	213	2	AAW14567 Streptoco
9	464	91.9	641	2	AAW61217 Streptoco
10	464	91.9	641	5	ABP54636 S. pneumo
11	464	91.9	641	7	ADC45241 S. pneumo
12	463	91.7	137	7	ABW02598 Ac122c pn
13	442	87.5	233	7	ABW02606 Bf1019c pn
14	439.5	87.0	196	2	AAW14564 Streptoco
15	435	86.1	233	2	AAW14572 Streptoco
16	298	59.0	211	7	ABW02621 Bg11703c
17	298	59.0	238	2	AAW14587 Streptoco
18	293	58.0	232	7	ABW02624 EF5688c p
19	293	58.0	275	8	ADOS2055 S. pneumo
20	293	58.0	369	8	ADK52496 alpha hel
21	293	58.0	458	2	AAW14592 Streptoco
22	293	58.0	458	7	ABW02626 EF5668 pn
23	293	58.0	633	8	ADK52495 PspA mole
24	293	58.0	653	8	ADOS2080 S. pneumo
25	283	56.0	212	2	AAW14588 Streptoco

26	283	56.0	212	7	ABW02622	Bg7817c p
27	282.5	55.9	233	2	AAW14590	Ado14590 Streptoco
28	253	50.1	459	8	ADOL5316	Ado15316 S. pneumo
29	240.5	47.6	184	2	AAW14589	Ado14589 Streptoco
30	236	46.7	185	7	ABW02623	Bg7561c p
31	178	35.2	487	8	ADR04321	Ado04321 Streptoco
32	178	35.2	489	8	ADOS2088	Ado52088 Streptoco
33	178	35.2	524	8	ADOS2082	E. coli B
34	178	35.2	627	8	ADOS2129	Ado52129 E. coli B
35	177.5	35.1	119	2	AAW46291	Ado46291 Pneumococ
36	177.5	35.1	215	7	ABW14563	Abw14563 Streptoco
37	177.5	35.1	215	7	ABW02597	Abw02597 Atcc6303c
38	165.5	32.8	290	8	ADOS2119	Ado52119 pYA3637 b
39	165.5	32.8	298	8	ADOS2127	Ado52127 pYA3637 b
40	160	31.7	230	8	ADOS2086	Ado52086 S. pneumo
41	160	31.7	230	8	ADR04319	Ado04319 Streptoco
42	126	25.0	550	8	ADK48356	Adk48356 Streptoco
43	126	25.0	550	8	ADR95223	Adr95223 Novel S.
44	126	25.0	623	6	ABU08494	Abu08494 Fragment
45	121.5	24.1	204	2	AAW14571	Ado14571 Streptoco

ALIGNMENTS

RESULT 1

ABW02601

ID ABW02601 standard; protein; 213 AA.

XX AC ABW02601;

XX DT 12-FEB-2004 (first entry)

XX DE Bg8090c pneumococcal surface protein A (PspA) central region.

XX DE Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;

XX KW immunological; gene therapy; immunostimulant.

XX OS Unidentified.

XX FH Key Location/Qualifiers

XX FT Misc-difference 2 /label= Unknown

XX US6592876-B1.

XX PN 15-JUL-2003.

XX PD 15-SEP-1995; 95US-00529055.

XX PR 20-APR-1993; 93US-00048896.

XX PR 06-JUN-1995; 95US-00465746.

XX PA (UABR-) UAB RES FOUND.

XX PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;

XX DR WPI; 2003-862841/80.

XX XX Immunological composition for obtaining expression products used for

XX PT detecting the presence of Streptococcus pneumoniae or its strain,

XX PT comprises at least two different full length isolated gene encoding

XX PT pneumococcal surface protein A.

XX XX Example 6; SEQ ID NO 47; 121pp; English.

XX CC The present invention relates to an immunological composition comprising at least 2 different full length isolated genes encoding pneumococcal surface protein A (PspA) from different groups based on restriction fragment polymorphism analysis. The invention is useful for obtaining expression products by recombinant techniques to detect, determine, isolate or diagnose the presence of Streptococcus pneumoniae or its strain. The expression product is useful for preparing antigenic,

CC immunological or vaccine compositions, for eliciting antibodies, an
 CC immunological response (other than or additional to antibodies) or a
 CC protective response (including antibody or other immunological response
 CC by administering compositions to a host). The invention is also useful as
 CC vaccines and in gene therapy. The present sequence is Bg8090c
 CC pneumococcal surface protein A (PspA) central region. This sequence is
 CC used in the exemplification of the invention
 XX
 SQ Sequence 213 AA;

Query Match 92.9%; Score 469; DB 7; Length 213;
 Best Local Similarity 95.2%; Pred. No. 2.4e-35;
 Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELKE-AEALDKKADLPNKVADLEKEISNLEILLG 58
 |||||
 DB 59 LAKKQTELEKLLDNLDPGKTQDELKEAEALDKKADLPNKVADLEKEISNLEILLG 118
 |||||

QY 59 GADSEDDTAALPNKLTAKAELEKTQKELDAALNELGPDGDEEE 102
 |||||
 DB 119 GADPEDDTAALPNKLTAKAEFEKTPKELDAALNELGPDGDEEE 162
 |||||

RESULT 2
 ABU08487
 ID ABU08487 standard; protein; 8991 AA.
 AC ABU08487;
 XX
 DT 24-JUN-2003 (first entry)
 XX
 DE S. pneumoniae pneumococcal surface protein A (PspA) protein.
 XX
 KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
 KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
 KW antibacterial.
 XX
 OS Streptococcus pneumoniae.
 XX
 PH Key Location/Qualifiers
 FT Misc-difference 1..8991
 FT /note= "All Xaa residues within this sequence are
 FT unknown"

XX US6500613-B1.
 XX
 XX 31-DEC-2002.
 XX
 XX 16-SEP-1996; 96US-00714741.
 XX
 XX 15-SEP-1995; 95US-00529055.
 XX
 XX (UYAL-) UNIV ALABAMA.
 XX
 XX Briles DE, Mcdaniel JS, Swiatlo E, Yother J, Crain MJ;
 PI Hollingshead S, Tart R, Brooks-Walter A;
 XX
 XX WPI; 2003-361534/34.
 XX
 XX Isolated PspC amino acid sequence used as polymerase chain reaction or
 PT hybridization probe, comprises pneumococcal surface protein having alpha-
 PT helical, proline rich and repeat regions.
 XX
 PS Disclosure; Col 145-188; 186pp; English.
 XX
 CC The present invention relates to the isolation of Streptococcus
 CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
 CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-
 CC like protein having alpha-helical, proline rich and repeat regions. The
 CC PspC and PspA proteins may be used in a vaccine to protect against
 CC pneumococcal infections. The polynucleotide sequences encoding PspC and
 CC PspA may be used for the expression of the proteins, and as PCR primers
 CC or hybridisation probes. The present sequence represents S. pneumoniae

CC PspA protein
 XX
 SQ Sequence 8991 AA;

Query Match 92.9%; Score 469; DB 6; Length 8991;
 Best Local Similarity 95.2%; Pred. No. 2e-33;
 Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELKE-AEALDKKADLPNKVADLEKEISNLEILLG 58
 |||||
 DB 4625 LAKKQTELEKLLDNLDPGKTQDELKEAEALDKKADLPNKVADLEKEISNLEILLG 4684
 |||||

QY 59 GADSEDDTAALPNKLTAKAELEKTQKELDAALNELGPDGDEEE 102
 |||||
 DB 4685 GADPEDDTAALPNKLTAKAEFEKTPKELDAALNELGPDGDEEE 4728
 |||||

RESULT 3
 ADK52498
 ID ADK52498 standard; protein; 416 AA.
 XX
 AC ADK52498;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE alpha helical region PspA molecule from the EF3296 strain.
 XX
 KW Streptococcus pneumoniae infection; pneumococcal surface protein A; PspA;
 KW hemolytic anemia disease; leukemia; lymphoma; sickle cell anemia;
 KW Hodgkin's disease.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO2004016231-A2.
 XX
 PD 26-FEB-2004.
 XX
 PF 17-FEB-2003; 2003WO-US008199.
 XX
 PR 15-MAR-2002; 2002US-0365351P.
 XX
 PA (UABR-) UAB RES FOUND.
 XX
 XX Briles DE;
 XX
 XX WPI; 2004-192068/18.
 XX
 PT Treating Streptococcus pneumoniae infection in a subject lacking a
 PT functional spleen comprises administering an antibody that recognizes
 PT pneumococcal surface protein A (PspA) or its binding portion.
 XX
 PS Claim 17; SEQ ID NO 4; 41pp; English.
 XX
 CC The present invention relates to treating Streptococcus pneumoniae
 CC infection in a subject lacking a functional spleen comprises
 CC administering an antibody that recognizes pneumococcal surface protein A
 CC (PspA) or its binding portion. The method is useful for treating or
 CC preventing Streptococcus pneumoniae infection in a subject lacking a
 CC functional spleen. The disease-associated injury is especially due to
 CC hemolytic anemia disease, leukemia or lymphoma, especially sickle cell
 CC anemia or Hodgkin's disease. The present sequence represents the alpha
 CC helical region PspA molecule from the EF3296 strain of Streptococcus
 CC pneumoniae.
 XX
 SQ Sequence 416 AA;

Query Match 92.7%; Score 468; DB 8; Length 416;
 Best Local Similarity 96.2%; Pred. No. 6.5e-35;
 Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELKEA-EALDKKADLPNKVADLEKEISNLEILLG 58
 |||||
 DB 241 LAKKQTELEKLLDNLDPGKTQDELKEAEALDKKADLPNKVADLEKEISNLEILLG 300
 |||||

QY 59 GADSEDDTAALPNKLTAKAELEKTKQKELDAALNELGPDGDEE 102
 |||||
 Db 301 GADSEDDTAALQNLKATKAELEKTKQKELDAALNELGPDGDEE 344

RESULT 4
 ADK52497
 ID ADK52497 standard; protein; 526 AA.

AC ADK52497;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE PspA molecule from the EF3296 strain of Streptococcus pneumoniae.

KW Streptococcus pneumoniae infection; pneumococcal surface protein A; PspA;
 KW hemolytic anemia disease; leukemia; lymphoma; sickle cell anemia;
 KW Hodgkin's disease.

XX Streptococcus pneumoniae.

OS WO2004016231-A2.

PN 26-FEB-2004.

PD 17-FEB-2003; 2003WO-US008199.

PF 15-MAR-2002; 2002US-0365351P.

PR (UABR-) UAB RES FOUND.

XX Briles DE;

XX WPI; 2004-192068/18.

XX Treating Streptococcus pneumoniae infection in a subject lacking a
 PT functional spleen comprises administering an antibody that recognizes
 PT pneumococcal surface protein A (PspA) or its binding portion.

XX Claim 17; SEQ ID NO 3; 41pp; English.

XX The present invention relates to treating Streptococcus pneumoniae
 CC infection in a subject lacking a functional spleen comprises
 CC administering an antibody that recognizes pneumococcal surface protein A
 CC (PspA) or its binding portion. The method is useful for treating or
 CC preventing Streptococcus pneumoniae infection in a subject lacking a
 CC functional spleen. The disease-associated injury is especially due to
 CC hemolytic anemia disease, leukemia or lymphoma, especially sickle cell
 CC anemia or Hodgkin's disease. The present sequence represents PspA
 CC molecule from the EF3296 strain of Streptococcus pneumoniae.

XX Sequence 526 AA;

Query Match 92.7%; Score 468; DB 8; Length 526;
 Best Local Similarity 96.2%; Pred. No. 8.6e-35;
 Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEKGTODELDKEA-EAELOKKADELQNKVADLEKEISNLEILG 58
 |||||
 Db 346 LAKKQTELEKLLDLSLQPEKGTODELDKEAEAELOKKADELQNKVADLEKEISNLEILG 405

QY 59 GADSEDDTAALPNKLTAKAELEKTKQKELDAALNELGPDGDEE 102
 |||||
 Db 406 GADSEDDTAALQNLKATKAELEKTKQKELDAALNELGPDGDEE 449

RESULT 5
 ABU00449
 ID ABU00449 standard; protein; 744 AA.

XX ABU00449;

XX

DT 23-OCT-2003 (revised)
 DT 11-FEB-2003 (first entry)
 XX
 DE S. pneumoniae type 4 strain protein from coding region #16.
 XX
 KW Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;
 KW antinflammatory; antibacterial; immunostimulant; auditory; respiratory;
 KW gene therapy; vaccine.
 XX
 OS Streptococcus pneumoniae; type 4 strain.
 XX
 PN WO200277021-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 27-MAR-2002; 2002WO-IB002163.
 XX
 PR 27-MAR-2001; 2001GB-00007658.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Masignani V, Tettelin H, Fraser C;
 XX
 DR WPI; 2003-040579/03.
 DR N-PSDB; ABX05728.
 XX
 PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,
 PT useful as medicaments for treating or preventing a disease or infection
 PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or
 PT ear infection.

Claim 1; SEQ ID NO 32; 56pp; English.

The invention relates to a protein comprising or having at least 50%
 identity to any of the 2469 amino acid sequences, identified in the
 specification (available on a computer readable format), or its fragment,
 expressed from 2469 identified DNA coding regions from the
 Streptococcus pneumoniae type 4 strain genomic sequence appearing as
 ABSS56454. Also included are an antibody which binds one of the proteins,
 treating a patient by administering the protein, DNA or antibody (in a
 composition), a kit comprising first and second primers, which are the
 nucleic acid cited above or fragments between nucleotides 8-100 of a
 sequence not defined in the specification, for amplifying a target
 sequence contained within a Streptococcus nucleic acid sequence, where
 the first primer is substantially complementary to the target sequence
 and the second primer is substantially complementary to the complement of
 the target sequence, and where the parts of the primers having
 substantial complementarity define the termini of the target sequence to
 be amplified, assay comprising contacting a test compound with the
 protein, and determining whether the test compound binds to the protein
 and a Streptococcus pneumoniae bacterium, where one or more genes
 encoding the proteins has been rendered inactive. The proteins, nucleic
 acid molecules, antibody and compositions are useful as medicaments for
 treating or preventing a disease or infection due to streptococcus
 bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis
 media or ear infection. They are also useful in developing vaccines,
 diagnostics and antibiotics. The methods are useful for identifying
 immunodominant proteins. The present sequence is one of the 2469 proteins
 expressed by the identified coding regions from the genomic sequence.
 CC Note: the sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to
 CC standardise OS field)

XX Sequence 744 AA;

Query Match 92.7%; Score 468; DB 6; Length 744;
 Best Local Similarity 96.2%; Pred. No. 1.3e-34;
 Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEKGTODELDKEA-EAELOKKADELQNKVADLEKEISNLEILG 58
 |||||

Db 346 LAKKQTELEKLLDLSLDPGKGTQDELKAEAEALDKKADLQNKVADLEKEISNLEILG 405

QY 59 GADSEDDTAALPNKLTAKKAELEKTQKELDAALNELGPDGDEE 102
 |||||

Db 406 GADSEDDTAALQNKLTAKKAELEKTQKELDAALNELGPDGDEE 449

RESULT 6
 ADM92054
 ID ADM92054 standard; protein; 744 AA.
 XX
 AC ADM92054;
 XX
 DT 03-JUN-2004 (first entry)
 XX
 DE S pneumoniae antigenic protein sequence SeqID251.
 XX
 KW antibacterial; gene therapy; Streptococcus pneumoniae infection;
 KW antigenic.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO2004020609-A2.
 XX
 PD 11-MAR-2004.
 XX
 PF 02-SEP-2003; 2003WO-US027401.
 XX
 PR 30-AUG-2002; 2002US-0407082P.
 XX
 PA (TUFT) UNIV TUFTS.
 XX
 PI Camilli A, Hava DL;
 XX
 DR WPI; 2004-239189/22.
 DR N-PSDB; ADM91817.
 XX
 PT New Streptococcus pneumoniae nucleic acid molecules, useful for
 PT diagnosing, treating and preventing active infections of Streptococcus
 PT pneumoniae.
 XX
 PS Claim 27; SEQ ID NO 251; 123pp; English.
 XX
 CC This invention relates to novel isolated Streptococcus pneumoniae nucleic
 CC acid molecules and the antigenic polypeptides encoded by them. The
 CC invention may be useful for the production of compounds with an
 CC antibacterial activity or for gene therapy. The nucleic acid molecules,
 CC compositions and methods disclosed are useful for treating Streptococcus
 CC pneumoniae infection. The present sequence is that of an S pneumoniae
 CC protein of the invention.
 XX
 SQ Sequence 744 AA;

Query Match 92.7%; Score 468; DB 8; Length 744;
 Best Local Similarity 96.2%; Pred. No. 1.3e-34;
 Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKGTQDELKAEAEALDKKADLQNKVADLEKEISNLEILG 58
 |||||

Db 346 LAKKQTELEKLLDLSLDPGKGTQDELKAEAEALDKKADLQNKVADLEKEISNLEILG 405

QY 59 GADSEDDTAALPNKLTAKKAELEKTQKELDAALNELGPDGDEE 102
 |||||

Db 406 GADSEDDTAALQNKLTAKKAELEKTQKELDAALNELGPDGDEE 449

RESULT 7
 AAY81652
 ID AAY81652 standard; protein; 745 AA.
 XX
 AC AAY81652;
 XX
 DT 24-MAY-2000 (first entry)

XX Streptococcus pneumoniae protein sequence ID301.
 DE
 XX
 KW Streptococcus pneumoniae; vaccine; screening; protein antigen;
 KW antibacterial; antiinflammatory; meningitis; infection; diagnosis;
 KW pneumococcal disease.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO200006737-A2.
 XX
 PD 10-FEB-2000.
 XX
 PF 27-JUL-1999; 99WO-GB002451.
 XX
 PR 27-JUL-1998; 98GB-00016337.
 PR 19-MAR-1999; 99US-0125164P.
 XX
 PA (MICR-) MICROBIAL TECHNIQS LTD.
 XX
 PI Gilbert CFG, Hansbro PM;
 XX
 DR WPI; 2000-195300/17.
 XX
 PT New Streptococcal protein, useful as a vaccine, for diagnosis of
 PT pneumococcal diseases and for screening agents capable of antagonizing or
 PT inhibiting expression of the protein.
 XX
 PS Claim 2; Page 95; 108pp; English.
 XX
 CC AAY81501 to AAY81679 represent specifically claimed protein sequences
 CC isolated from Streptococcus pneumoniae. AAA05407 to AAA05590 represent
 CC specifically claimed nucleotide sequences isolated from S. pneumoniae.
 CC The sequences have antibacterial and antiinflammatory properties. The
 CC protein sequences, and fragments of them, are useful as immunogens and/or
 CC antigens. The nucleotide sequences can be used in vaccines and in
 CC diagnostic assays. The proteins and nucleotides can be useful for the
 CC detection and diagnosis of S. pneumoniae. The protein sequences are also
 CC useful for screening an agent capable of antagonising, inhibiting or
 CC interfering with the function or expression of the proteins in which the
 CC agent is useful for treatment or prophylaxis of S. pneumoniae infection
 CC and meningitis. AAA05591 to AAA05614 represent primers used in the
 CC exemplification of the present invention
 XX
 SQ Sequence 745 AA;

Query Match 92.7%; Score 468; DB 3; Length 745;
 Best Local Similarity 96.2%; Pred. No. 1.3e-34;
 Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKGTQDELKAEAEALDKKADLQNKVADLEKEISNLEILG 58
 |||||

Db 346 LAKKQTELEKLLDLSLDPGKGTQDELKAEAEALDKKADLQNKVADLEKEISNLEILG 405

QY 59 GADSEDDTAALPNKLTAKKAELEKTQKELDAALNELGPDGDEE 102
 |||||

Db 406 GADSEDDTAALQNKLTAKKAELEKTQKELDAALNELGPDGDEE 449

RESULT 8
 AAW14567
 ID AAW14567 standard; protein; 213 AA.
 XX
 AC AAW14567;
 XX
 DT 17-OCT-2003 (revised)
 DT 28-OCT-1997 (first entry)
 XX
 DE Streptococcus pneumoniae PspA central region.
 XX
 KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
 KW bacteraemia; pneumonia.
 XX

OS	Streptococcus pneumoniae; strain Bg8090.					
XX	Key	Location/Qualifiers				
XX	Misc-difference 2	/note= "unidentified amino acid"				
FT						
FT						
XX	PN	W09709994-A1.				
XX	PD	20-MAR-1997.				
XX	XX					
XX	PP	16-SEP-1996; 96WO-US014819.				
XX	XX					
PR	PR	15-SEP-1995; 95US-00529055.				
XX	PA	(UABR-) UAB RES FOUND.				
XX	PI	Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;				
XX	PI	Hollingshead S, Tart R, Brooks-Walter A;				
XX	XX					
DR	WT	1997-202002/18.				
XX	XX					
PT	PT	Streptococcus pneumoniae surface protein PspC and truncated PspA - used				
PT	PT	in vaccines for protecting animals against S.pneumoniae infection.				
XX	PS	Example 6; Fig 13; 296pp; English.				
XX	CC					
XX	CC	This sequence shows the central portion, including the C-terminus of the				
CC	CC	alpha-helix region and some of the proline-rich region, of pneumococcal				
CC	CC	surface protein A (PspA) of Streptococcus pneumoniae strain Bg8090.				
CC	CC	Comparison of the N-terminal and central regions (AAWL4533-57 and				
CC	CC	AAWL4562-91) of PspA polypeptides from different pneumococcal strains can				
CC	CC	be used to divide the strains into several families based on sequence				
CC	CC	homologies. PspA polypeptides, or fragments of them, can be used in				
CC	CC	vaccines to protect animals against S. pneumoniae infection and hence for				
CC	CC	the prevention of diseases such as otitis media, meningitis, bacteraemia				
CC	CC	and pneumonia. The sequence of the 3' half of the PspA alpha-helical				
CC	CC	region and the immediate 5' tip of the coding sequence are likely to be				
CC	CC	the critical sequences for predicting PspA cross-reactions and vaccine				
CC	CC	composition. (Updated on 17-OCT-2003 to standardise OS field)				
XX	XX					
SQ	SQ	Sequence 213 AA;				
Query Match		92.5%; Score 467; DB 2; Length 213;				
Best Local Similarity		95.2%; Pred. No. 3.7e-35;				
Matches		99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;				
QY	1	LAKKQTLEKLKD-LDPEGKTQDELDKE-AAEALDKKADELPNKVADLKEISNLEILLG	58			
Dd	59	LAKKQTLEKLKDNLDPGKTKQDELDKEAAEAELDKKADELPNKVADLKEISNLEILLG	118			
QY	59	GADSEDDTAALPNKLTAKAELEKTOKELDAALNELGPDGDDEE	102			
Dd	119	GADPEDDTAALPNKLTAKAEFEKTFKELDAALNELGPDGDDEE	162			
RESULT 9						
AAW61217						
ID	AAW61217	standard; protein; 641 AA.				
AC	AAW61217;					
XX						
DT	02-OCT-1998	(first entry)				
XX						
DE	Streptococcus pneumoniae SP0092 protein.					
XX						
KW	Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;					
KW	detection; pneumonia; otitis media; meningitis.					
XX						
OS	Streptococcus pneumoniae...					
XX						
FH	Key	Location/Qualifiers				
FT	Misc-difference 306	/label= unknown				
FT						

FT		/note= "encoded by NCT"	
XX	WO9818930-A2.		
PN	XX		
PD	XX		
PP	07-MAY-1998.		
Pf	30-OCT-1997; 97WO-US019422.		
XX	XX		
PR	31-OCT-1996; 96US-0029960P.		
XX	XX		
PA	(HUMA-) HUMAN GENOME SCI INC.		
PI	Kunsch CA, Choi GH, Johnson LS, Hromockyj A;		
XX	XX		
DR	WPI; 1998-272224/24.		
DR	N-PSDB; AAV27403.		
XX	XX		
PT	Nucleic acid encoding antigenic peptide(s) from Streptococcus pneumoniae		
PT	- or their epitope-containing fragments, useful in protective or		
PT	therapeutic vaccines, and for diagnosis.		
XX	XX		
PS	Claim 11; Page 82; 118pp; English.		
XX	XX		
CC	The present sequence represents a protein from Streptococcus pneumoniae.		
CC	The nucleic acid sequence encoding the Streptococcus pneumoniae protein		
CC	can be useful in vaccines for inducing protective antibodies against		
CC	Streptococcus pneumoniae, for treatment or prevention of infection e.g.		
CC	pneumonia, otitis media or meningitis. Probes based on the nucleic acid		
CC	are used to detect Streptococcus infection (by usual hybridisation or		
CC	amplification methods), also for isolating Streptococcus genes or their		
CC	allelic variants. The protein can be used similarly to detect specific		
CC	antibodies in standard immunoassays, especially for diagnosing or		
CC	monitoring infections. Antibodies which bind the protein are used to		
CC	detect corresponding antigens, to purify the protein and for passive		
CC	immunisation (optionally coupled to a toxin). Vaccines are administered,		
CC	e.g. by injection, orally or through the skin, typically at 0.01-1000		
CC	(especially 10-300) mu g/ml per dose		
XX	XX		
SQ	Sequence 641 AA;		
	Query Match 91.9%; Score 464; DB 2; Length 641;		
	Best Local Similarity 95.2%; Pred. NO. 2.5e-34;		
	Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;		
QY	1 LAKKQTELEKLDD-LDPEGTODELDKEA-EAEELDKADELPKNVADLEKEISNLEILLG 58		
Dd	243 LAKKQTELEKLDDSDPEGTODELDKEAEAEELDKADELPKNVADLEKEISNLEILLG 302		
QY	59 GADSEDDTAALPNKLATKKAELEKTOKELDAALNELPGPDDEE 102		
Dd	303 GADXEDDTAALQNKLATKKAELEKTOKELDAALNELPGPDDEE 346		
RESULT 10			
ABP54636	ID		
AC	ABP54636 standard; protein; 641 AA.		
AC	ABP54636;		
XX	XX		
XX	04-SEP-2002 (first entry)		
XX	XX		
DE	S. pneumoniae SP092 protein sequence SEQ ID NO:160.		
XX	Streptococcus pneumoniae; epitope; vaccine; antigenic protein;		
KW	antibacterial; Streptococcal infection; detection.		
XX	XX		
OS	Streptococcus pneumoniae.		
XX	XX		
PN	US2002061545-A1.		
XX	XX		
PD	23-MAY-2002.		
XX	XX		
PF	22-JAN-2001; 2001US-00765272.		

```

XX 30-OCT-1997; 97US-00961083.
XX (CHOI/) CHOI G H.
XX (KUNS/) KUNSCH C A.
XX (BARA/) BARASH S C.
XX (DILL/) DILLON P J.
XX (DOUG/) DOUGHERTY B.
XX (FANN/) FANNON M R.
XX (ROSE/) ROSEN C A.
XX
XX Choi GH, Kunsch CA, Barash SC, Dillon PJ, Dougherty B, Fannon MR;
XX Rosen CA;
XX WPI: 2003-764574/72.
XX N-PSDB; ABC45240.
XX
XX Novel polynucleotide encoding Streptococcus pneumoniae polypeptides
XX useful for producing vaccines for prevention or attenuation of infection
XX by Streptococcus pneumoniae.
XX
XX Example 1; SEQ ID NO 160; 50pp; English.
XX
XX The invention relates to an isolated polynucleotide consisting of a
XX Streptococcus pneumoniae nucleic acid (appearing as ADC45122 and encoding
XX SP028) one of 113 disclosed nucleic acids encoding 113 S. pneumoniae
XX antigens. Also included are making a recombinant vector by inserting the
XX nucleic acid into a vector, an isolated polynucleotide consisting of at
XX least 50 or 100 contiguous nucleotides of the SP028 nucleic acid, and a
XX recombinant host cell comprising the SP028 polynucleotide. The nucleic
XX acids are useful as DNA vaccine against Streptococcus pneumoniae
XX infection (e.g. pneumonia). Nucleic acids derived from the S. pneumoniae
XX antigen nucleic acids are useful as probes for use in diagnostic methods
XX for detecting S. pneumoniae gene expression. The present sequence
XX represents an S. pneumoniae antigenic protein.
XX
XX SQ Sequence 641 AA;
XX
XX Query Match 91.9%; Score 464; DB 5; Length 641;
XX Best Local Similarity 95.2%; Pred. No. 2.5e-34;
XX Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
XX
XX 1 LAKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPKNKADLEKEISNLEILLG 58
XX 243 LAKQTELEKLLDLSLDPEGKTQDELKKEAEAEALDKKADLPKNKADLEKEISNLEILLG 302
XX
XX 59 GADSEDDTAALPNKATKKAELKTKQKELDAALNELGPDGDEEE 102
XX 303 GADXEDDTAALQNKLATKKAELKTKQKELDAALNELGPDGDEEE 346
XX
XX RESULT 11
XX ADC45241
XX ID ADC45241 standard; protein; 641 AA.
XX
XX AC ADC45241;
XX
XX 18-DEC-2003 (first entry)
XX
XX S. pneumoniae antigenic protein SP092.
XX
XX Antigen; bacterial infection; vaccine; pneumonia; antibacterial.
XX
XX Streptococcus pneumoniae.
XX
XX US6573082-B1.
XX
XX 03-JUN-2003.
XX
XX 28-MAR-2000; 2000US-00536784.
XX
XX 31-OCT-1996; 96US-0029960P.
XX
XX 30-OCT-1997; 97US-00961083.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX

```

```

XX Choi GH, Kunsch CA, Barash SC, Dillon PJ, Dougherty B, Fannon MR;
XX Rosen CA;
XX WPI: 2003-764574/72.
XX N-PSDB; ADC45240.
XX
XX Novel polynucleotide encoding Streptococcus pneumoniae polypeptides
XX useful for producing vaccines for prevention or attenuation of infection
XX by Streptococcus pneumoniae.
XX
XX Example 1; SEQ ID NO 160; 50pp; English.
XX
XX The invention relates to an isolated polynucleotide consisting of a
XX Streptococcus pneumoniae nucleic acid (appearing as ADC45122 and encoding
XX SP028) one of 113 disclosed nucleic acids encoding 113 S. pneumoniae
XX antigens. Also included are making a recombinant vector by inserting the
XX nucleic acid into a vector, an isolated polynucleotide consisting of at
XX least 50 or 100 contiguous nucleotides of the SP028 nucleic acid, and a
XX recombinant host cell comprising the SP028 polynucleotide. The nucleic
XX acids are useful as DNA vaccine against Streptococcus pneumoniae
XX infection (e.g. pneumonia). Nucleic acids derived from the S. pneumoniae
XX antigen nucleic acids are useful as probes for use in diagnostic methods
XX for detecting S. pneumoniae gene expression. The present sequence
XX represents an S. pneumoniae antigenic protein.
XX
XX SQ Sequence 641 AA;
XX
XX Query Match 91.9%; Score 464; DB 7; Length 641;
XX Best Local Similarity 95.2%; Pred. No. 2.5e-34;
XX Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
XX
XX 1 LAKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPKNKADLEKEISNLEILLG 58
XX 243 LAKQTELEKLLDLSLDPEGKTQDELKKEAEAEALDKKADLPKNKADLEKEISNLEILLG 302
XX
XX 59 GADSEDDTAALPNKATKKAELKTKQKELDAALNELGPDGDEEE 102
XX 303 GADXEDDTAALQNKLATKKAELKTKQKELDAALNELGPDGDEEE 346
XX
XX RESULT 12
XX ABW02598
XX ID ABW02598 standard; protein; 197 AA.
XX
XX AC ABW02598;
XX
XX 12-FEB-2004 (first entry)
XX
XX Ac122c pneumococcal surface protein A (PspA) central region.
XX
XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
XX immunological; gene therapy; immunostimulant.
XX
XX Unidentified.
XX
XX US6592876-B1.
XX
XX 15-JUL-2003.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX 20-APR-1993; 93US-00048896.
XX
XX 06-JUN-1995; 95US-00465746.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX WPI: 2003-862841/80.
XX
XX Immunological composition for obtaining expression products used for
XX detecting the presence of Streptococcus pneumoniae or its strain,
XX comprises at least two different full length isolated gene encoding
XX

```

```

PT pneumococcal surface protein A.
PS Example 6; SEQ ID NO 44; 121pp; English.
XX
CC The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspAs) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antibodies, an
CC immunological or vaccine compositions, for eliciting antigenic, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Acl22c pneumococcal
CC surface protein A (PspA) central region. This sequence is used in the
CC exemplification of the invention
XX
SQ Sequence 197 AA;

Query Match      91.7%; Score 463; DB 7; Length 197;
Best Local Similarity 94.2%; Pred. No. 7.8e-35;
Matches 98; Conservative 2; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLD-LDPEGKTQDELDKA-EALDKKADLPNKVADLKEISNLEILG 58
DB 22 LAQKQTELGKLLDLDPEGKTQDELDKAAGEALDKKADGLPNKVSDLEKEISNLEILG 81
QY 59 GADSEDDTAALPNKLTAKKAELEKTKQELDAALNELGPDGDEE 102
DB 82 GADSEDDTAALPNKLTAKKAELEKTKQELDAALNELGPDGDEE 125

RESULT 13
ABW02606
ID ABW02606 standard; protein; 233 AA.
XX
AC ABW02606;
XX
DT 12-FEB-2004 (first entry)
XX
DE Bfl019c pneumococcal surface protein A (PspA) central region.
XX
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
KW immunological; gene therapy; immunostimulant.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..233
FT /note= "Xaa = Unknown amino acid"
XX
PN US6592876-B1.
XX
PD 15-JUL-2003.
XX
PP 15-SEP-1995; 95US-00529055.
XX
PR 20-APR-1993; 93US-00048896.
PR 06-JUN-1995; 95US-00465746.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX
DR WPI; 2003-862841/80.
XX
CC Immunological composition for obtaining expression products used for
CC detecting the presence of Streptococcus pneumoniae or its strain.
CC comprises at least two different full length isolated gene encoding
CC pneumococcal surface protein A.
XX

```

```

PS Example 6; SEQ ID NO 52; 121pp; English.
XX
CC The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspAs) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antibodies, an
CC immunological or vaccine compositions, for eliciting antibodies, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Efl019c
CC pneumococcal surface protein A (PspA) central region. This sequence is
CC used in the exemplification of the invention
XX
SQ Sequence 233 AA;

Query Match      87.5%; Score 442; DB 7; Length 233;
Best Local Similarity 91.3%; Pred. No. 8.3e-33;
Matches 95; Conservative 1; Mismatches 6; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLD-LDPEGKTQDELDKA-EALDKKADLPNKVADLKEISNLEILG 58
DB 53 LAQKQTELGKLLDLDPEGKTQDELDKAAGEALDKKADLPNKVADLKEISNLEILG 112
QY 59 GADSEDDTAALPNKLTAKKAELEKTKQELDAALNELGPDGDEE 102
DB 113 GADSEDDTAALPNKLTAKKAELEKTKQELDAALNELGPDGDEE 156

RESULT 14
AAW14564
ID AAW14564 standard; protein; 196 AA.
XX
AC AAW14564;
XX
DT 17-OCT-2003 (revised)
DT 28-OCT-1997 (first entry)
XX
DE Streptococcus pneumoniae PspA central region.
XX
KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW bacteraemia; pneumonia.
XX
OS Streptococcus pneumoniae; strain Acl22.
XX
PN WO9709994-A1.
XX
PD 20-MAR-1997.
XX
PP 16-SEP-1996; 96WO-US014819.
XX
PR 15-SEP-1995; 95US-00529055.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX
DR WPI; 1997-202002/18.
XX
PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used
PT in vaccines for protecting animals against S.pneumoniae infection.
XX
PS Example 6; Fig 13; 296pp; English.
XX
CC This sequence shows the central portion, including the C-terminus of the
CC alpha-helix region and some of the proline-rich region, of pneumococcal
CC surface protein A (PspA) of Streptococcus pneumoniae strain Acl22.
CC Comparison of the N-terminal and central regions (AAW14533-57 and
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
XX

```

be used to divide the strains into several families based on sequence homologies. PspA polypeptides, or fragments of them, can be used in CC vaccines to protect animals against *S. pneumoniae* infection and hence for CC the prevention of diseases such as otitis media, meningitis, bacteraemia CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical CC region and the immediate 5' tip of the coding sequence are likely to be CC the critical sequences for predicting PspA cross-reactions and vaccine CC composition. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 196 AA;
SQ

Query Match	87.0%;	Score 439.5;	DB 2;	Length 196;
Best Local Similarity	92.3%;	Pred. No. 1.1e-32;		
Matches	96;	Conservative 2;	Mismatches 3;	Gaps 3;

QY	1	LAKQTELEKLLD-LDPEGKTQDLDKEA-EAELDKKADELPNKVADLEKESINLEILLG	58
		:	
Db	22	LAQQTGLKLLDLPDKTKTQDLDKEAGEALDKKADGLPNKVSQLEKESINLEILLG	81
		:	
QY	59	GAISEDTPALPNKLTATKKALEKTKQELDAAINELPGDGEDE	102
Db	82	GAISEDTPALPNKLTATKKALEKTKQELDAAINELPGDGEDE	124

RESULT 15
AAW14572
ID AAW14572 standard; protein; 233 AA.

AC	AAW14572;
XX	
XX	17-OCT-2003 (revised)
DT	28-OCT-1997 (first entry)
DT	
XX	
XX	Streptococcus pneumoniae PspA central region.
DE	
XX	
XX	PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW	bacteraemia; pneumonia.
KW	
XX	
XX	Streptococcus pneumoniae; strain Ef3296.
OS	

Key	Location/Qualifiers
FT	Misc-difference 129
FT	/note= "unidentified amino acid"
FT	Misc-difference 131
FT	/note= "unidentified amino acid"
XX	
PN	WO9709994-A1.

XX	
PD	
XX	20-MAR-1997.
XX	
XX	
PF	16-SEP-1996; 96WO-US014819.
XX	
XX	
PR	15-SEP-1995; 95US-00529055.
XX	
PA	(UABR-) UAB RES FOUND.
XX	
XX	Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain WJ;
PI	Hollingshead S, Tart R, Brooks-Walter A;
XX	
XX	WPI; 1997-202002/18.
DR	

Streptococcus pneumoniae surface protein PspC and truncated PspA - used in vaccines for protecting animals against S.pneumoniae infection.

PS Example 6; Fig 13; 296pp; English.

This sequence shows the central portion, including the C-terminus of the alpha-helix region and some of the proline-rich region, of pneumococcal surface protein A (PspA) of *Streptococcus pneumoniae* strain EF3286. Comparison of the N-terminal and central regions (AA114533-57 and AA114562-91) of PspA polypeptides from different pneumococcal strains can be used to divide the strains into several families based on sequence homologies. PspA polypeptides, or fragments of them, can be used in

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:55:14 ; Search time 19.0469 Seconds
(without alignments)
399.760 Million cell updates/sec

Title: US-10-674-755-18

Perfect score: 505

Sequence: 1 LAKQTELEKLLDPEGKT.....TQKELDALNELGPDGDEE 102

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/prodata/1/iaa/PCUTS_COMB.pep.*
- 6: /cgn2_6/prodata/1/iaa/backfilese1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	505	100.0	102	2	US-08-710-749-21
2	505	100.0	102	4	US-09-147-875A-18
3	484	95.8	104	2	US-08-710-749-19
4	484	95.8	104	4	US-09-147-875A-20
5	479	94.9	104	4	US-09-147-875A-21
6	473	93.7	104	2	US-08-710-749-20
7	469	92.9	213	4	US-08-529-055-47
8	469	92.9	8991	4	US-08-714-741-32
9	464	91.9	641	3	US-08-561-083-160
10	464	91.9	641	4	US-09-536-784-160
11	463	91.7	137	4	US-08-529-055-44
12	442	87.5	233	4	US-08-529-055-52
13	369.5	73.2	80	2	US-08-710-749-18
14	369.5	73.2	80	4	US-09-147-875A-19
15	298	59.0	108	2	US-08-710-749-26
16	298	59.0	108	4	US-09-147-875A-23
17	298	59.0	211	4	US-08-529-055-67
18	293	58.0	108	2	US-08-710-749-24
19	293	58.0	108	4	US-09-147-875A-25
20	293	58.0	232	4	US-08-529-055-70
21	293	58.0	458	4	US-08-529-055-73
22	290	57.4	108	4	US-09-147-875A-24
23	288	57.0	106	4	US-09-147-875A-22
24	283	56.0	212	4	US-08-529-055-68
25	280	55.4	108	2	US-08-710-749-22
26	280	55.4	108	2	US-08-710-749-23
27	243	48.1	108	2	US-08-710-749-25

28 243 48.1 108 4 US-09-147-875A-26 Sequence 26, Appl
29 236 46.7 185 4 US-08-529-055-69 Sequence 69, Appl
30 177.5 35.1 119 2 US-08-710-749-27 Sequence 27, Appl
31 177.5 35.1 119 4 US-09-147-875A-27 Sequence 27, Appl
32 177.5 35.1 215 4 US-08-529-055-43 Sequence 43, Appl
33 126 25.0 550 4 US-09-583-110-4871 Sequence 4871, Ap
34 126 25.0 550 4 US-09-107-433-3858 Sequence 3858, Ap
35 126 25.0 623 4 US-08-714-741-47 Sequence 47, Appl
36 121.5 24.1 101 2 US-08-710-749-1 Sequence 1, Appl
37 121.5 24.1 204 4 US-08-529-055-51 Sequence 51, Appl
38 121 24.0 605 4 US-08-714-741-46 Sequence 46, Appl
39 120.5 23.9 101 2 US-08-710-749-2 Sequence 2, Appl
40 119.5 23.7 289 1 US-08-072-070-4 Sequence 4, Appl
41 119.5 23.7 289 1 US-08-469-434-4 Sequence 4, Appl
42 119.5 23.7 289 1 US-08-214-222-4 Sequence 4, Appl
43 119.5 23.7 289 2 US-08-467-852A-5 Sequence 5, Appl
44 119.5 23.7 289 2 US-08-468-718-4 Sequence 4, Appl
45 119.5 23.7 289 2 US-08-247-491A-5 Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-710-749-21
; Sequence 21, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Brilles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 102 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-21

Query Match 100.0%; Score 505; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 2.7e-42;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LAKQTELEKLLDPEGKTQDELKAEAEFLDKKADLPNKVADLEKEISNLEILGGA 60
DB 1 LAKQTELEKLLDPEGKTQDELKAEAEFLDKKADLPNKVADLEKEISNLEILGGA 60

```
Qy 61 DSEDDTAALPNKLTAKAELEKTKQKELDAALNELGPDGDEEE 102
      |||||
Db 61 DSEDDTAALPNKLTAKAELEKTKQKELDAALNELGPDGDEEE 102

RESULT 2
US-09-147-875A-18
; Sequence 18, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-18

Query Match 100.0%; Score 505; DB 4; Length 102;
Best Local Similarity 100.0%; Pred. No. 2.7e-42;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LAKQTELEKLLDLPDGKTKQDELDKAEAEALDKKADLPNKVADLEKEISNLEILGGA 60
      |||||
Db 1 LAKQTELEKLLDLPDGKTKQDELDKAEAEALDKKADLPNKVADLEKEISNLEILGGA 60

Qy 61 DSEDDTAALPNKLTAKAELEKTKQKELDAALNELGPDGDEEE 102
      |||||
Db 61 DSEDDTAALPNKLTAKAELEKTKQKELDAALNELGPDGDEEE 102

RESULT 3
US-08-710-749-19
; Sequence 19, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Brilles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 104 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-19

Query Match 95.8%; Score 484; DB 2; Length 104;
Best Local Similarity 98.1%; Pred. No. 3.1e-40;
Matches 102; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 1 LAKQTELEKLLD-LDPEGKTKQDELDKAEAEALDKKADLPNKVADLEKEISNLEILG 58
      |||||
Db 1 LAKQTELEKLLD-LDPEGKTKQDELDKAEAEALDKKADLPNKVADLEKEISNLEILG 60

Qy 59 GADSEDDTAALPNKLTAKAELEKTKQKELDAALNELGPDGDEEE 102
      |||||
Db 61 GADSEDDTAALPNKLTAKAELEKTKQKELDAALNELGPDGDEEE 104
```

```
RESULT 4
US-09-147-875A-20
; Sequence 20, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-20
```

```
Query Match 95.8%; Score 484; DB 4; Length 104;
Best Local Similarity 98.1%; Pred. No. 3.1e-40;
Matches 102; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 1 LAKQTELEKLLD-LDPEGKTKQDELDKAEAEALDKKADLPNKVADLEKEISNLEILG 58
      |||||
Db 1 LAKQTELEKLLD-LDPEGKTKQDELDKAEAEALDKKADLPNKVADLEKEISNLEILG 60

Qy 59 GADSEDDTAALPNKLTAKAELEKTKQKELDAALNELGPDGDEEE 102
      |||||
Db 61 GADSEDDTAALPNKLTAKAELEKTKQKELDAALNELGPDGDEEE 104
```

```
RESULT 5
US-09-147-875A-21
; Sequence 21, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-21
```

```
Query Match 94.9%; Score 479; DB 4; Length 104;
Best Local Similarity 97.1%; Pred. No. 9.6e-40;
Matches 101; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
```



```
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8991 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-714-741-32

Query Match 92.9%; Score 469; DB 4; Length 8991;
Best Local Similarity 95.2%; Pred. No. 2.3e-36;
Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELDKA-EAELDKKADELPNKVADLEKEISNLEILLG 58
DB 4625 LAKKQTELEKLLDLDPEGKTQDELDKAEEAELDKKADELPNKVADLEKEISNLEILLG 4684

QY 59 GADSEDDTAALPNKLTAKAELEKTKQELDAALNELGPDGDEE 102
DB 4685 GADPEDDTAALPNKLTAKAELEKTKQELDAALNELGPDGDEE 4728

RESULT 9
US-08-961-083-160
; Sequence 160, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB340P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 160:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 160:
US-08-961-083-160

Query Match 91.9%; Score 464; DB 4; Length 641;
Best Local Similarity 95.2%; Pred. No. 2.7e-37;
Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELDKA-EAELDKKADELPNKVADLEKEISNLEILLG 58
DB 243 LAKKQTELEKLLDLDPEGKTQDELDKAEEAELDKKADELPNKVADLEKEISNLEILLG 302

QY 59 GADSEDDTAALPNKLTAKAELEKTKQELDAALNELGPDGDEE 102
DB 303 GADXEDDTAALQNKLTAKAELEKTKQELDAALNELGPDGDEE 346

RESULT 10
US-09-536-784-160
; Sequence 160, Application US/09536784
; Patent No. 6573082
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/536,784
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB340P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 160:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 160:
US-09-536-784-160

Query Match 91.9%; Score 464; DB 4; Length 641;
Best Local Similarity 95.2%; Pred. No. 2.7e-37;
Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELDKA-EAELDKKADELPNKVADLEKEISNLEILLG 58
```



```
|||||
Db 243 LAKQTELEKLDLSDPEGKTQDELKAEAEELDKKADLQNKVADLEKEISNLEILG 302
QY 59 GADSEDDTAALPNKLATKKALEKTKQKELDAALNELGPDGDEE 102
Db 303 GADXEDDTAALQNKLATKKALEKTKQKELDAALNELGPDGDEE 346

RESULT 11
US-08-529-055-44
; Sequence 44, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yotter, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 197 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-44

Query Match 91.7%; Score 463; DB 4; Length 197;
Best Local Similarity 94.2%; Pred. No. 7.7e-38;
Matches 98; Conservative 2; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKQTELEKLD-LDPEGKTQDELKAEAEELDKKADLQNKVADLEKEISNLEILG 59
Db 22 LAKQTELEKLDLSDPEGKTQDELKAEAEELDKKADLQNKVADLEKEISNLEILG 81

QY 59 GADSEDDTAALPNKLATKKALEKTKQKELDAALNELGPDGDEE 102
Db 82 GADSEDDTAALPNKLATKKALEKTKQKELDAALNELGPDGDEE 125

RESULT 12
US-08-529-055-52
; Sequence 52, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
```

```
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yotter, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 233 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-52

Query Match 87.5%; Score 442; DB 4; Length 233;
Best Local Similarity 91.3%; Pred. No. 1.1e-35;
Matches 95; Conservative 1; Mismatches 6; Indels 2; Gaps 2;

QY 1 LAKQTELEKLD-LDPEGKTQDELKAEAEELDKKADLQNKVADLEKEISNLEILG 58
Db 53 LAKQTELEKLDLSDPEGKTQDELKAEAEELDKKADLQNKVADLEKEISNLEILG 112

QY 59 GADSEDDTAALPNKLATKKALEKTKQKELDAALNELGPDGDEE 102
Db 113 GADSEDDTAALPNKLATKKALEKTKQKELDAALNELGPDGDEE 156

RESULT 13
US-08-710-749-18
; Sequence 18, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
```

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/710,749
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-710-749-18

Query Match 73.2%; Score 369.5; DB 2; Length 80;
Best Local Similarity 96.2%; Pred. No. 3.5e-29;
Matches 77; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 24 LDKEA-EAEIDKKADELPNKVADLEKEISNLEILLGADSEDDTAALPNKLTAKAELEK 82
Db 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGADSEDDTAALPNKLTAKAELEK 60

Qy 83 TOKELDAALNELGPDGDEEE 102
Db 61 TOKELDAALNELGPDGDEEE 80

RESULT 14
US-09-147-875A-19
; Sequence 19, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-147-875A-19

Query Match 73.2%; Score 369.5; DB 4; Length 80;
Best Local Similarity 96.2%; Pred. No. 3.5e-29;
Matches 77; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 24 LDKEA-EAEIDKKADELPNKVADLEKEISNLEILLGADSEDDTAALPNKLTAKAELEK 82
Db 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGADSEDDTAALPNKLTAKAELEK 60

Qy 83 TOKELDAALNELGPDGDEEE 102
Db 61 TOKELDAALNELGPDGDEEE 80

RESULT 15
US-08-710-749-26
; Sequence 26, Application US/08710749
; Patent No. 5955089
```

```
;
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 108 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-710-749-26

Query Match 59.0%; Score 298; DB 2; Length 108;
Best Local Similarity 64.8%; Pred. No. 5e-22;
Matches 70; Conservative 12; Mismatches 20; Indels 6; Gaps 4;

Qy 1 LAKKQTELEKLLD-LDPEGKTQDELKKE-AEAEIDKKADELPNKVADLEKEISNLEILLG 58
Db 1 LEXAAELENLLSTLDPEGKTQDELKKEAAEAEIKNKKVEALPNQVSELEBELSKLEDNLK 60

Qy 59 GADS---ED-DTAALPNKLTAKAELEKTKQKELDAALNELGPDGDEEE 102
Db 61 DAEITNNVEDYIKEGLEEAATKQAELEKTPKELDAALNELGPDGDEEE 108
```

Search completed: June 21, 2005, 10:25:22
Job time : 19.0469 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 10:12:15 ; Search time 65.1551 Seconds
(without alignments)
601.118 Million cell updates/sec

Title: US-10-674-755-18

Perfect score: 505

Sequence: 1 LAKKQTELEKLLDPEGKT.....TKELDAAALNELGPDGDEEE 102

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

```

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	505	100.0	102	15	US-10-674-755-18
2	484	95.8	104	15	US-10-674-755-20
3	479	94.9	104	15	US-10-674-755-21
4	469	92.9	213	15	US-10-299-636-62
5	468	92.7	744	10	US-09-769-787-184
6	468	92.7	744	17	US-10-472-928-32
7	464	91.9	641	9	US-09-765-272-160
8	463	91.7	197	15	US-10-299-636-59
9	442	87.5	233	15	US-10-299-636-67
10	369.5	73.2	80	15	US-10-674-755-19
11	298	59.0	108	15	US-10-674-755-23

12	298	59.0	211	15	US-10-299-636-82	Sequence 82, Appl
13	293	58.0	108	15	US-10-674-755-25	Sequence 25, Appl
14	293	58.0	232	15	US-10-299-636-85	Sequence 85, Appl
15	293	58.0	275	16	US-10-414-532-1	Sequence 1, Appl
16	293	58.0	458	15	US-10-299-636-88	Sequence 88, Appl
17	293	58.0	653	16	US-10-414-532-26	Sequence 26, Appl
18	290	57.4	108	15	US-10-674-755-24	Sequence 24, Appl
19	288	57.0	106	15	US-10-674-755-22	Sequence 22, Appl
20	283	56.0	212	15	US-10-299-636-83	Sequence 83, Appl
21	253	50.1	459	16	US-10-702-305A-18	Sequence 18, Appl
22	243	48.1	108	15	US-10-674-755-26	Sequence 26, Appl
23	236	46.7	185	15	US-10-299-636-84	Sequence 84, Appl
24	178	35.2	487	16	US-10-414-532-34	Sequence 34, Appl
25	178	35.2	487	16	US-10-414-533-21	Sequence 21, Appl
26	178	35.2	524	16	US-10-414-532-28	Sequence 28, Appl
27	177.5	35.1	119	15	US-10-674-755-27	Sequence 27, Appl
28	177.5	35.1	215	15	US-10-299-636-58	Sequence 58, Appl
29	165.5	32.8	290	16	US-10-414-532-65	Sequence 65, Appl
30	160	31.7	230	16	US-10-414-532-32	Sequence 32, Appl
31	160	31.7	230	16	US-10-414-533-19	Sequence 19, Appl
32	121.5	24.1	204	15	US-10-299-636-66	Sequence 66, Appl
33	119	23.6	354	15	US-10-299-636-105	Sequence 105, Appl
34	119	23.6	588	15	US-10-299-636-96	Sequence 96, Appl
35	119	23.6	619	10	US-09-882-774-1	Sequence 1, Appl
36	119	23.6	619	15	US-10-282-122A-73702	Sequence 73702, A
37	119	23.6	619	16	US-10-414-532-72	Sequence 72, Appl
38	118	23.4	141	14	US-10-254-995-2	Sequence 2, Appl
39	118	23.4	589	9	US-09-748-875-14	Sequence 14, Appl
40	118	23.4	589	10	US-09-298-523B-14	Sequence 14, Appl
41	118	23.4	589	15	US-10-299-636-97	Sequence 97, Appl
42	118	23.4	643	15	US-10-299-636-95	Sequence 95, Appl
43	118	23.4	670	9	US-09-748-875-63	Sequence 63, Appl
44	118	23.4	670	10	US-09-298-523B-63	Sequence 63, Appl
45	118	23.4	690	9	US-09-748-875-61	Sequence 61, Appl

ALIGNMENTS

RESULT 1

```

US-10-674-755-18
; Sequence 18, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-10-674-755-18

```

Query Match 100.0%; Score 505; DB 15; Length 102;
Best Local Similarity 100.0%; Pred. No. 7.8e-36;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	LAKKQTELEKLLDPEGKTQDELKAEAEALDKKADLPNKVADLEKISNLEILGGA	60
Db	1	LAKKQTELEKLLDPEGKTQDELKAEAEALDKKADLPNKVADLEKISNLEILGGA	60
QY	61	DSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE	102
Db	61	DSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE	102

RESULT 2

```
US-10-674-755-20
; Sequence 20, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
;
US-10-674-755-20
Query Match 95.8%; Score 484; DB 15; Length 104;
Best Local Similarity 98.1%; Pred. No. 5e-34;
Matches 102; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
Qy 1 LAKKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPNPKVADLEKEISNLEILLG 58
Db 1 LAKKQTELEKLLDLDPEGKTQDELKKEA-EAELDKKADLPNPKVADLEKEISNLEILLG 60
Qy 59 GADSEDDTAALPNKLTAKKAELEKTQKELDAALNELGPDGDEE 102
Db 61 GAUSDDTAALPNKLTAKKAELEKTQKELDAALNELGPDGDEE 104
RESULT 3
US-10-674-755-21
; Sequence 21, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
;
US-10-674-755-21
Query Match 94.9%; Score 479; DB 15; Length 104;
Best Local Similarity 97.1%; Pred. No. 1.3e-33;
Matches 101; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
Qy 1 LAKKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPNPKVADLEKEISNLEILLG 58
Db 1 LAKKQTELEKLLDLDPEGKTQDELKKEA-EAELDKKADLPNPKVADLEKEISNLEILLG 60
Qy 59 GADSEDDTAALPNKLTAKKAELEKTQKELDAALNELGPDGDEE 102
Db 61 GADPEDDTAALPNKLTAKKAELEKTQKELDAALNELGPDGDEE 104
RESULT 4
US-10-299-636-62
; Sequence 62, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
```

```
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (2)
; OTHER INFORMATION: Xaa at position 2 is unknown
;
US-10-299-636-62
Query Match 92.9%; Score 469; DB 15; Length 213;
Best Local Similarity 95.2%; Pred. No. 2.2e-32;
Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
Qy 1 LAKKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPNPKVADLEKEISNLEILLG 58
Db 59 LAKKQTELEKLLDLDPEGKTQDELKKEA-EAELDKKADLPNPKVADLEKEISNLEILLG 118
Qy 59 GADSEDDTAALPNKLTAKKAELEKTQKELDAALNELGPDGDEE 102
Db 119 GADPEDDTAALPNKLTAKKAELEKTQKELDAALNELGPDGDEE 162
RESULT 5
US-09-769-787-184
; Sequence 184, Application US/09769787
; Publication No. US20030091577A1
; GENERAL INFORMATION:
; APPLICANT: Microbial
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 184
; LENGTH: 744
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
;
US-09-769-787-184
Query Match 92.7%; Score 468; DB 10; Length 744;
Best Local Similarity 96.2%; Pred. No. 1.2e-31;
Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;
Qy 1 LAKKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPNPKVADLEKEISNLEILLG 58
Db 346 LAKKQTELEKLLDLDPEGKTQDELKKEA-EAELDKKADLPNPKVADLEKEISNLEILLG 405
Qy 59 GADSEDDTAALPNKLTAKKAELEKTQKELDAALNELGPDGDEE 102
Db 406 GADSEDDTAALPNKLTAKKAELEKTQKELDAALNELGPDGDEE 449
```

RESULT 6
US-10-472-928-32
; Sequence 32, Application US/10472928
; Publication No. US20050020813A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926W0
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: Seqwin99, version 1.03
; SEQ ID NO 32
; LENGTH: 744
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: pneumococcal surface protein A (pspA)
; OTHER INFORMATION: Cellular location: outside
; OTHER INFORMATION: Feature of note: WYV motif
; OTHER INFORMATION: Similar to strain R6 sequence 15902165 (e-179)
US-10-472-928-32

Query Match 92.7%; Score 468; DB 17; Length 744;
Best Local Similarity 96.2%; Pred. No. 1.2e-31;
Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELDEKA-EAELDKKADLPNKVADLEKEISNLEILG 58
|||||
Db 346 LAKKQTELEKLLDSDPEGKTQDELDEKA-EAELDKKADLPNKVADLEKEISNLEILG 405
|||||

QY 59 GADSEDDTAALPNKLTAKKAELEKTQKELDAALNELGPDGDEE 102
|||||
Db 406 GADSEDDTAALQNKLTAKKAELEKTQKELDAALNELGPDGDEE 449
|||||

RESULT 7
US-09-765-272-160
; Sequence 160, Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 160:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 160:
US-09-765-272-160

Query Match 91.9%; Score 464; DB 9; Length 641;
Best Local Similarity 95.2%; Pred. No. 2.1e-31;
Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELDEKA-EAELDKKADLPNKVADLEKEISNLEILG 58
|||||
Db 243 LAKKQTELEKLLDSDPEGKTQDELDEKA-EAELDKKADLPNKVADLEKEISNLEILG 302
|||||

QY 59 GADSEDDTAALPNKLTAKKAELEKTQKELDAALNELGPDGDEE 102
|||||
Db 303 GADXEDDTAALQNKLTAKKAELEKTQKELDAALNELGPDGDEE 346
|||||

RESULT 8
US-10-299-636-59
; Sequence 59, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-59

Query Match 91.7%; Score 463; DB 15; Length 197;
Best Local Similarity 94.2%; Pred. No. 6.6e-32;
Matches 98; Conservative 2; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELDEKA-EAELDKKADLPNKVADLEKEISNLEILG 58
|||||
Db 22 LAQKQTELEKLLDSDPEGKTQDELDEKA-EAELDKKADLPNKVADLEKEISNLEILG 81
|||||

QY 59 GADSEDDTAALPNKLTAKKAELEKTQKELDAALNELGPDGDEE 102
|||||
Db 82 GADSEDDTAALPNKLTAKKAELEKTQKELDAALNELGPDGDEE 125
|||||

RESULT 9
US-10-299-636-67
; Sequence 67, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E

QY 1 LAKQTELEKLL-DLDPEGKTQDELDK-EEAELDKADELPNKVADLEKEISNLEILLG 58

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:53:39 ; Search time 10.2 Seconds
(without alignments)
962.168 Million cell updates/sec

Title: US-10-674-755-18
Perfect score: 505
Sequence: 1 LAKKQTELEKLLDLPDGKT.....TKELDAALNELGPDGDEEE 102
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	468	92.7	744	2 F95013	pneumococcal surfa
2	119	23.6	619	2 A97887	surface protein ps
3	119	23.6	619	2 A41971	surface protein ps
4	106.5	21.1	924	2 S06117	myosin heavy chain
5	106.5	21.1	2007	1 B43402	myosin heavy chain
6	105	20.8	629	2 T44607	hypothetical prote
7	102	20.2	281	2 F75216	hypothetical prote
8	100	19.8	233	2 S70531	bbk2.11 protein pr
9	99.5	19.7	284	2 S23470	beta-tropomyosin -
10	99.5	19.7	1976	2 A59252	myosin heavy chain
11	99	19.6	415	2 S35760	fcra protein precu
12	98.5	19.5	388	2 A46173	Mrp4 protein - Str
13	98	19.4	284	2 A45488	body-wall muscle t
14	97.5	19.3	770	2 S56805	probable RNA helic
15	97	19.2	281	2 A34787	tropomyosin 1 alph
16	97	19.2	284	1 TMRBA	tropomyosin alpha
17	97	19.2	284	2 A39816	tropomyosin 2, fib
18	97	19.2	284	2 B27407	tropomyosin 2, fib
19	97	19.2	284	2 A27674	tropomyosin 3, fib
20	97	19.2	284	2 A25925	tropomyosin 3, fib
21	97	19.2	284	2 A60597	tropomyosin 2, fib
22	97	19.2	284	2 B39816	tropomyosin 3, fib
23	97	19.2	516	2 B84709	hypothetical prote
24	97	19.2	864	2 B90395	purine NTPase [imp
25	97	19.2	2139	2 T18296	myosin heavy chain
26	96.5	19.1	397	2 H86754	prophage pi2 prote
27	96.5	19.1	1932	2 A47297	myosin heavy chain
28	96	19.0	284	2 S24972	tropomyosin alpha,
29	96	19.0	284	2 S19691	tropomyosin alpha,

30	96	19.0	1063	2 T18255	cytoskeleton assem
31	96	19.0	1837	2 T41023	probable nuclear p
32	96	19.0	1972	1 A41604	myosin heavy chain
33	95.5	18.9	388	2 S52536	fcra 15 protein -
34	95.5	18.9	405	2 A33939	Fc gamma (IgG) rec
35	95.5	18.9	1509	1 A27224	myosin heavy chain
36	95	18.8	280	2 A22165	tropomyosin alpha
37	95	18.8	1169	2 A64505	P115 homolog - Met
38	95	18.8	1177	2 B75150	chromosome segrega
39	94.5	18.7	473	2 F70031	cell wall-binding
40	94.5	18.7	880	2 F75103	conserved hypothet
41	94	18.6	308	2 T08796	tropomyosin - huma
42	94	18.6	385	2 T20410	hypothetical prote
43	93.5	18.5	1175	2 D35815	myosin heavy chain
44	93.5	18.5	1201	2 B35815	myosin heavy chain
45	93.5	18.5	1475	2 T33318	hypothetical prote

ALIGNMENTS

RESULT 1

F95013
pneumococcal surface protein A [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C:Accession: F95013
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzap
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: F95013
A>Status: preliminary
A:Molecule type: DNA
A;Residues: 1-744 <KUR>
A;Cross-references: UNIPROT:Q97T39; GB:AE005672; PIDN:AAK74303.1; PID:gl4971584; GSPDB:G
A;Experimental source: strain TIGR4
C:Genetics:
A;Gene: SP0117

Query Match 92.7%; Score 468; DB 2; Length 744;
Best Local Similarity 96.2%; Pred. No. 2.1e-25;
Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY	1	LAKKQTELEKLLD-LDPEGKTQDELDKA-EAELDKKADLQNKVADLEKEISNLEILLG	58
DB	346	LAKKQTELEKLLDLSLDPEGKTQDELDKAEAEELDKKADLQNKVADLEKEISNLEILLG	405
QY	59	GADSEDDTAAALPNKATKKALEKTKQKELDAALNELGPDGDEEE	102
DB	406	GADSEDDTAAALQNKLATKKALEKTKQKELDAALNELGPDGDEEE	449

RESULT 2

A97887
surface protein pepA precursor [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; B
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: A97887
A>Status: preliminary
A:Molecule type: DNA
A;Residues: 1-619 <KUR>

F:1316-2007/Region: light meromyosin
F:129/Modified site: N6,N6-crimethyllsine (lys) #status predicted
F:184/Binding site: ATP (Lys) #status predicted
F:732/742/Active site: Cys #status predicted
F:1954/Binding site: phosphate (Thr) (covalent) #status predicted
F:1987/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 21.1%; Score 106.5; DB 1; Length 2007;
Best Local Similarity 31.5%; Pred. No. 8.2;
Matches 34; Conservative 25; Mismatches 42; Indels 7; Gaps 3;

QY 1 LAKKQTELEKLLD-----LDPEKTDDELKAEAELEKQKELDAALNELPGDDEE 102
Db 1056 LAKKNQEMWITDLERLKEETKQELK-ARRKLDGETTDLQQAELQAIQELKI 1114
QY 56 LIGGADSEDDTA-ALPNKATKKALEKTKELDAALNELPGDDEE 102
Db 1115 QLAKEEELQAALARGDEEAVQKNALKVIRELQAQIAELQEDLESEK 1162

RESULT 6
T44607
hypothetical protein hp71 - Halobacterium salinarum
C:Species: Halobacterium salinarum
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T44607
R:Ruepp, A.; Wanner, G.; Soppa, J.
Arch. Microbiol. 169, 1-9, 1998
A:Title: A 71-kDa protein from Halobacterium salinarum belongs to a ubiquitous P-loop A
A:Reference number: Z22810; MUID:98060711; PMID:9396829
A:Accession: T44607
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-629 <RUE>
A:Cross-references: UNIPROT:O07116; EMBL:Y13615; PIDN:CRA73936.1
A:Note: the source is designated as Halobacterium salinarum
C:Genetics:
A:Note: hp71
C:Function:
A:Description: might be involved in cytoskeleton formation and/or chromosome partitionin

Query Match 20.8%; Score 105; DB 2; Length 629;
Best Local Similarity 33.7%; Pred. No. 3.2;
Matches 35; Conservative 22; Mismatches 23; Indels 24; Gaps 6;

QY 3 KKQTELE-KLLDLDPGKTDDELKAEAELEKKA--DELPNKVADLEKEISNLEILGG 59
Db 177 QKVTRLQGEHLDN---ETFDELERKVDSETDQRLDELTAQSOLEREAEARLE----- 228
QY 60 ADSEDDTAALPNKATKKALEKTKELDAALNELG-PGDDEE 102
Db 229 -----RLENQVDRRKQALD----DKEAALESLDIPDSPTAE 260

RESULT 7
F75216
hypothetical protein PAB2181 - Pyrococcus abyssi (strain Oresay)
C:Species: Pyrococcus abyssi
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: F75216
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A:Reference number: A75001
A:Accession: F75216
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-281 <KAW>
A:Cross-references: UNIPROT:Q9V217; GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB4918
A:Experimental source: strain Oresay
C:Genetics:
A:Gene: PAB2181

Query Match 20.2%; Score 102; DB 2; Length 281;
Best Local Similarity 31.6%; Pred. No. 2.2;
Matches 31; Conservative 20; Mismatches 18; Indels 18; Gaps 4;

QY 1 LAKKQTELE-KLLDLDPGKTDDELKAEAELEKKADELPNKVADLEKEISNLEIL 56
Db 182 LEKLLKESEVKLMYEAKAKRAEELKLEKVEEKV-KREESLERKVSERSLSNEYTK 240
QY 57 LGGADSEDDTAALPNKATKKALEKTKELDAALNEL 94
Db 241 V-----KSLEKKKELEKNKVELEEVNKL 265

RESULT 8
S70531
bbk2.11 protein precursor - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C>Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S70531
R:Akins, D.R.; Porcella, S.F.; Popova, T.G.; Shevchenko, D.; Baker, S.I.; Li, M.; Norgard
Mol. Microbiol. 18, 507-520, 1995
A:Title: Evidence for in vivo but not in vitro expression of a Borrelia burgdorferi outer
A:Reference number: S70531; MUID:96342380; PMID:8748034
A:Accession: S70531
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-233 <AKI>
A:Cross-references: UNIPROT:Q44739; EMBL:U30617; NID:g3309515; PIDN:AAC46421.1; PID:g119;
C:Superfamily: outer surface protein F ospF
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-233/Product: bbk2.11 protein #status predicted <MAT>

Query Match 19.8%; Score 100; DB 2; Length 233;
Best Local Similarity 30.8%; Pred. No. 2.5;
Matches 33; Conservative 16; Mismatches 38; Indels 20; Gaps 3;

QY 3 KKQTE--LEKLLDLDPGKTDDELKAEAELEKKADELPNKV-----ADLE 47
Db 51 KKQVEGFLLE----ETKDLKDEKOTKEIKQELKNKIEKLDKSKTSIETSYE 105
QY 48 KEISNLEILGGADSEDDTAALPNKATKKALEKTKELDAALNEL 94
Db 106 EKINKIKEKLGKLEDEKFELESLESLAKKGERKALQEAQKFEEL 152

RESULT 9
S23470
beta-tropomyosin - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S23470
R:Hardy, S.; Thiebaud, P.
Biochim. Biophys. Acta 1131, 239-242, 1992
A:Title: Isolation and characterization of cDNA clones encoding the skeletal and smooth
A:Reference number: S23470; MUID:92305070; PMID:1610908
A:Accession: S23470
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-284 <HAR>
A:Cross-references: UNIPROT:Q91706; EMBL:M87307; NID:g214001; PIDN:AAA18100.1; PID:g2140
A:Note: the authors translated the codon GAG for residue 14 as Asp
C:Superfamily: tropomyosin

Query Match 19.7%; Score 99.5; DB 2; Length 284;
Best Local Similarity 32.6%; Pred. No. 3.4;
Matches 31; Conservative 21; Mismatches 32; Indels 11; Gaps 4;

QY 3 KKQTE-----LEKLLDLDPGK-TODELDKAEAELE--KKADELPNKVADLEKEISNL 53
Db 29 KKQAEKRCQKIEBEVMDLQKTKGTDEVEKYSESVKDAQEKLEMAEKATDAEAVASL 88
QY 54 EILGGADSEDDTAALPNKATKKALEKTKELDKQLD 88

Db 89 NRRIQVVEELDRA--QERLATALQKLEETEKAVID 121

RESULT 10

A59252

myosin heavy chain, nonmuscle, form IIB - human

N;Alternate names: myosin type 10; NMHC-B

N;Contains: MYOSIN ATPase (EC 3.6.4.1)

C;Species: Homo sapiens (man)

C;Date: 19-May-2000 #sequence revision 19-May-2000 #text_change 09-Jul-2004

C;Accession: A59252; B61231; G02055

R;Phillips, C.L.; Yamakawa, K.; Adelstein, R.S.

J. Muscle Res. Cell. Motil. 16, 379-389, 1995

A;Title: Cloning of the cDNA encoding human nonmuscle myosin heavy chain-B and analysis

A;Reference number: A59252; MUID:96025307; PMID:7499478

A;Accession: A59252

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-1976 <SIM>

A;Cross-references: UNIPROT:P35580; GB:M69181; NID:G641957; PIDN:AAA99177.1; PID:G641958

A;Experimental source: clone lib Lambda Zap II adult human T-cell library; cell line Jur

A;Note: between nucleotides 1942-1943 in mRNA encoding human brain MHC-B there is an alt

R;Simons, M.; Wang, M.; McBride, O.W.; Kawamoto, S.; Yamakawa, K.; Gdula, D.; Adelstein, R.

Circ. Res. 69, 530-539, 1991

A;Title: Human nonmuscle myosin heavy chains are encoded by two genes located on differe

A;Reference number: A61231; MUID:91316803; PMID:1860190

A;Accession: B61231

A;Molecule type: mRNA

A;Residues: 63-237, 'K', 239-664, 'L', 666-722 <SI2>

A;Cross-references: GB:M69181; NID:G641957

R;Weir, L.

submitted to the EMBL Data Library, August 1995

A;Reference number: H00753

A;Accession: G02055

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-81 <WEI>

A;Cross-references: EMBL:U34304; NID:g1143217; PIDN:AAA84880.1; PID:g1143218

C;Genetics:

A;Gene: GDB:MVH10

A;Cross-references: GDB:127350; GDB:G00-127-350; OMIM:160776

A;Map position: 17p13-17p13

C;Superfamily: myosin heavy chain; myosin motor domain homology

C;Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; nucleotid

F;88-771/Domain: myosin motor domain homology <NMO>

F;178-185/Region: nucleotide-binding motif A (P-loop)

F;559-572/Region: actin binding #status predicted

F;633-647/Region: actin binding #status predicted

F;129/Modified site: N6,N6-trimethyllysine (Lys) #status predicted

F;184/Binding site: ATP (Lys) #status predicted

F;701,711/Active site: Cys #status predicted

Query Match 19.7%; Score 99.5; DB 2; Length 1976;

Best Local Similarity 28.4%; Pred. No. 25;

Matches 33; Conservative 24; Mismatches 36; Indels 23; Gaps 4;

Qy 1 LAKQTELEKLD-----LDPEGTQDELKAEAELEKKADELPNKVADLEKEISNLEI 55

Db 1025 LAKIRNQEWISDLERLKKKEETROLEK-ARKKLDGTTQDQIAELQAIQDELK 1083

Qy 56 LLGGADSE-----DDTAALPNKLTATKAELEKTKELDAALNELGPDGDEE 102

Db 1084 QLAKEEELQALARGDDETLHKNAL-----KVVRELQAIQAELEQDFESEK 1131

RESULT 11

S35760

fcrA protein precursor - Streptococcus pyogenes

C;Species: Streptococcus pyogenes

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004

C;Accession: S35760; A42711

R;Podbielski, A.

submitted to the EMBL Data Library, November 1992

A;Reference number: S35760

A;Accession: S35760

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-415 <FOD>

A;Cross-references: UNIPROT:Q54859; EMBL:X69324; NID:g311759; PIDN:CAA49165.1; PID:g3117

R;Haanes, E.J.; Heath, D.G.; Cleary, P.P.

J. Bacteriol. 174, 4967-4976, 1992

A;Title: Architecture of the vir regulons of group A streptococci parallels opacity fact

A;Reference number: A42711; MUID:92332431; PMID:1385809

A;Accession: A42711

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 343-415 <HAA>

A;Cross-references: GB:M86806; NID:g153630; PIDN:AAA26887.1; PID:g153631

A;Experimental source: strain CS101, OF+

A;Note: sequence extracted from NCBI backbone (NCBIN:108942, NCBI:P:108945)

C;Superfamily: M5 protein

Query Match 19.6%; Score 99; DB 2; Length 415;

Best Local Similarity 30.2%; Pred. No. 5.4;

Matches 39; Conservative 15; Mismatches 47; Indels 28; Gaps 4;

Qy 1 LAKQTELEKLDLDPEGTQDELKAEAE-----ELDKKADELPNKVAD 45

Db 188 LEAKNAEIEDLKQD-ASKTEETIANLQSEAAITLENLLGSAKHELTDLQAKLDTATAEKAK 246

Qy 46 LEKEISNLEILGGADSE-----DDTAALPNKLTATKAELEK---TOKELDAALNE 93

Db 247 LESQETTLNLLGSAKRELTDLQAKLDDANAEEKLQSQAALEKQLEATKKELADLOAK 306

Qy 94 LGPDGDEE 102

Db 307 LAATNQEKE 315

RESULT 12

A46173

Mrp4 protein - Streptococcus sp. (group A)

C;Species: Streptococcus sp.

C;Date: 21-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 30-May-1997

C;Accession: A46173

R;O'Toole, P.; Stenberg, L.; Rissler, M.; Lindahl, G.

Proc. Natl. Acad. Sci. U.S.A. 89, 8661-8665, 1992

A;Title: Two major classes in the M protein family in group A streptococci.

A;Reference number: A46173; MUID:92409576; PMID:1528877

A;Contents: group A

A;Accession: A46173

A;Status: preliminary

A;Molecule type: nucleic acid

A;Residues: 1-388 <OLT>

A;Note: sequence extracted from NCBI backbone (NCBIN:114063, NCBI:P:114064)

C;Superfamily: M5 protein

Query Match 19.5%; Score 98.5; DB 2; Length 388;

Best Local Similarity 31.6%; Pred. No. 5.4;

Matches 37; Conservative 15; Mismatches 40; Indels 25; Gaps 4;

Qy 1 LAKQTE-----LEKLLDLDPEGTQDELKAEAELEKKADELPNKVADLEKEISNLEIL 57

Db 182 IAKLQSEAAITLENLL-----GSAKREL-----TELOAKLDTATAEKAKLESQVTTLENLL 231

Qy 58 GGADSE-----DDTAALPNKLTATKAELEKTKELDAALNELGPDGDEE 102

Db 232 GSARELTDLQAKLDAANAEEKLQSQATLEKQLEATKKELADLOAKLAATNQEKE 288

RESULT 13

A45488

body-wall muscle tropomyosin - sea squirt (Ciona intestinalis)

C;Species: Ciona intestinalis

C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C;Accession: A45488

C:Accession: A34787
 E:Lees-Miller, J.P.; Goodwin, L.O.; Helfman, D.M.
 Mol. Cell. Biol. 10, 1729-1742, 1990
 A:Title: Three novel brain tropomyosin isoforms are expressed from the rat alpha-tropomyosin
 A:Reference number: A34787; MUID:90205854; PMID:2320008
 A:Accession: A34787
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-281 <LEE>
 A:Cross-references: UNIPROT:P04692; GB:M34135
 C:Superfamily: tropomyosin
 C:Keywords: alternative splicing; coiled coil

```
Query Match      19.2%; Score 97; DB 2; Length 281;
Best Local Similarity 27.5%; Pred. No. 5;
Matches 30; Conservative 20; Mismatches 31; Indels 28; Gaps 3;

Qy   10 KLIDLDPGKTDQLDKEAEALDKKADE-----LPNKVADLEKETSNLILLG----- 58
      :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db   9 QMLKLDK-----NALDRAEQAEAKKAEDRSQKQDELVLSQLKLKGTELDKYSEAL 64

Qy   59 -----GASDDTAAIPNKLATTKAELKQTOKELDAAINEL 94

Db   65 KDAQEKLEIAEKKATAEADVASLNRIOLIVBELDRAGERLATLQKL 113
```

Search completed: June 21, 2005, 10:12:04
Job time : 11.2 secs

```

Query Match      19.3%; Score 97.5; DB 2; Length 770;
Best Local Similarity 31.1%; Pred. No. 13;
Matches 33; Conservative 14; Mismatches 50; Indels 9; Gaps 3;

Qy 1 LAKQTELEK-----LLLDLDPGGKTQDELDKAAEALDKKADELPKNVADLEKEISNLEIL 56
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 597 LSKKASLASKGNASKLIFDDEGEAPVVELEDEEFPHKRGDAEVQKTEFLTKE-----SAV 652

Qy 57 LGGADSDDDTAALPNKLATKKAELKTKQKELDAAALNELGPPGDDEE 102
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 653 MADIDNIDKQVAKKKQKCKKRLKLEAMRREMEAAEER-EIRGSDDEE 697

```

RESULT 15
A34787

tropomyosin 1 alpha, brain - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 13-Jul-1990 #sequence revision 13-Jul-1990 #text change 09-Jul-2004

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:54:24 ; Search time 63.1776 Seconds
(without alignments)
826.751 Million cell updates/sec

Title: US-10-674-755-18
Perfect score: 505
Sequence: 1 LAKQTELEKLDLPDGKT.....TQKELDAALNELGPDGDEE 102

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 03: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	468	92.7	228	2	Q9L5B8	Q9L5B8 streptococc
2	468	92.7	235	2	Q9L582	Q9L582 streptococc
3	468	92.7	249	2	Q9L5D4	Q9L5D4 streptococc
4	468	92.7	252	2	Q9L583	Q9L583 streptococc
5	468	92.7	360	2	Q8KQK3	Q8KQK3 streptococc
6	468	92.7	429	2	Q9LAX7	Q9LAX7 streptococc
7	468	92.7	526	2	Q9LAX9	Q9LAX9 streptococc
8	468	92.7	608	2	Q8VQ55	Q8VQ55 streptococc
9	468	92.7	744	2	Q97T39	Q97T39 streptococc
10	463	91.7	231	2	Q9L579	Q9L579 streptococc
11	463	91.7	241	2	Q9L580	Q9L580 streptococc
12	463	91.7	249	2	Q9L5B7	Q9L5B7 streptococc
13	459	90.9	502	2	Q9LAX8	Q9LAX8 streptococc
14	457	90.5	249	2	Q9L585	Q9L585 streptococc
15	457	90.5	256	2	Q9L590	Q9L590 streptococc
16	454	89.9	242	2	Q9L562	Q9L562 streptococc
17	450	89.1	209	2	Q9L593	Q9L593 streptococc
18	432	85.5	222	2	Q9L584	Q9L584 streptococc
19	303	60.0	246	2	Q9L5B4	Q9L5B4 streptococc
20	299	59.2	479	2	Q9LAX2	Q9LAX2 streptococc
21	299	59.2	481	2	Q9LAX5	Q9LAX5 streptococc
22	298	59.0	107	2	Q8KQK2	Q8KQK2 streptococc
23	293	58.0	653	2	Q34Q97	Q34Q97 streptococc
24	291	57.6	213	2	Q8GNS7	Q8GNS7 streptococc
25	274	54.3	480	2	Q9LAX3	Q9LAX3 streptococc
26	236.5	46.8	211	2	Q8GNT0	Q8GNT0 streptococc
27	236.5	46.8	227	2	Q9L594	Q9L594 streptococc
28	230.5	45.6	227	2	Q9KGS0	Q9KGS0 streptococc
29	230.5	45.6	256	2	Q9L595	Q9L595 streptococc
30	230.5	45.6	461	2	Q9LAX6	Q9LAX6 streptococc
31	134	26.5	417	2	Q9LAY3	Q9LAY3 streptococc

ALIGNMENTS

RESULT 1

Q9L5B8 PRELIMINARY; PRT; 228 AA.
AC Q9L5B8;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=60;
RX MEDLINE=20472598; PubMed=11015380; Hollingshead S.K.; Beall B., Gherardi G., Facklam R.R.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant pneumococcal strains in the United States and of internationally disseminated clones."
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=60;
RX Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF253404; AAF67352.1;
DR InterPro; IPR009053; Prefoldin.
FT NON_TER 1
FT NON_TER 228
SQ SEQUENCE 228 AA; 24430 MW; E6EAA953EC54EA0F CRC64;

Query Match 92.7%; Score 468; DB 2; Length 228;
Best Local Similarity 96.2%; Pred. No. 4.9e-24;
Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;
QY 1 LAKQTELEKLD-LDPEKTDQELDKEA-EAELEKKADELKPKVADLEKEISNLEILLG 58
Db 46 LAKQTELEKLDLDPEKTDQELDKEAEEAELEKKADELKPKVADLEKEISNLEILLG 105
QY 59 GADSEDDTAAALPNKATKAELEKTKQELDKAALNELGPDGDEE 102
Db 106 GADSEDDTAAALQNKATKAELEKTKQELDKAALNELGPDGDEE 149

RESULT 2

Q9L582 PRELIMINARY; PRT; 235 AA.
ID Q9L582;
AC Q9L582;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE PspA (Fragment).

```

GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Packlam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RN J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=3;
RX Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255545; AAF68098.1; -.
FT NON_TER 1
FT NON_TER 235
SQ SEQUENCE 235 AA; 25424 MW; BFFB848C52CA8380 CRC64;

Query Match 92.7%; Score 468; DB 2; Length 235;
Best Local Similarity 96.2%; Pred. No. 5e-24;
Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Qy 1 LAKKQTELEKLLD-LDPEGKTQDELKKA-EAELDKKADELQKNKVADELEKEISNLEILLG 58
Db 56 LAKKQTELEKLLDLDPEGTQDELKKA-EAELDKKADELQKNKVADELEKEISNLEILLG 115
Qy 59 GADSEDDTAALPNKLTATKKAELKTKQKELDAALNELGPDGDEE 102
Db 116 GAUSSEDDTAALQNKLTATKKAELKTKQKELDAALNELGPDGDEE 159

RESULT 3
Qy 1 LAKKQTELEKLLD-LDPEGKTQDELKKA-EAELDKKADELQKNKVADELEKEISNLEILLG 58
Db 56 LAKKQTELEKLLDLDPEGTQDELKKA-EAELDKKADELQKNKVADELEKEISNLEILLG 115
Qy 59 GADSEDDTAALPNKLTATKKAELKTKQKELDAALNELGPDGDEE 102
Db 116 GAUSSEDDTAALQNKLTATKKAELKTKQKELDAALNELGPDGDEE 159

ID Q9L5D4 PRELIMINARY; PRT; 249 AA.
AC Q9L5D4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPI95;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Packlam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RN J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SPI95;
RX Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF252286; AAF69499.1; -.
FT NON_TER 1
FT NON_TER 249
SQ SEQUENCE 249 AA; 26821 MW; F8EA39225CF8D43F CRC64;

Query Match 92.7%; Score 468; DB 2; Length 249;
Best Local Similarity 96.2%; Pred. No. 5.3e-24;
Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

```

```

Qy 1 LAKKQTELEKLLD-LDPEGKTQDELKKA-EAELDKKADELQKNKVADELEKEISNLEILLG 58
Db 67 LAKKQTELEKLLDLDPEGTQDELKKA-EAELDKKADELQKNKVADELEKEISNLEILLG 126
Qy 59 GADSEDDTAALPNKLTATKKAELKTKQKELDAALNELGPDGDEE 102
Db 127 GADSEDDTAALQNKLTATKKAELKTKQKELDAALNELGPDGDEE 170

RESULT 4
ID Q9L583 PRELIMINARY; PRT; 252 AA.
AC Q9L583;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=127;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Packlam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RN J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=127;
RX Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255544; AAF68097.1; -.
FT NON_TER 1
FT NON_TER 252
SQ SEQUENCE 252 AA; 27260 MW; 82DE13741F369CA2 CRC64;

Query Match 92.7%; Score 468; DB 2; Length 252;
Best Local Similarity 96.2%; Pred. No. 5.4e-24;
Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Qy 1 LAKKQTELEKLLD-LDPEGKTQDELKKA-EAELDKKADELQKNKVADELEKEISNLEILLG 58
Db 73 LAKKQTELEKLLDLDPEGTQDELKKA-EAELDKKADELQKNKVADELEKEISNLEILLG 132
Qy 59 GADSEDDTAALPNKLTATKKAELKTKQKELDAALNELGPDGDEE 102
Db 133 GADSEDDTAALQNKLTATKKAELKTKQKELDAALNELGPDGDEE 176

RESULT 5
ID Q8KQK3 PRELIMINARY; PRT; 360 AA.
AC Q8KQK3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=259/98;
RX MEDLINE=22170754; PubMed=12183557;

```



```

RX DOI=10.1128/IAI.70.9.5086-5090.2002;
RA Miyaji E.N., Ferreira D.M., Lopes A.P.Y., Brandileone M.C.C.,
RA Dias W.O., Leite L.C.C.;
RT "Analysis of serum cross-reactivity and cross-protection elicited by
RT immunization with DNA vaccines against Streptococcus pneumoniae
RT expressing PspA fragments from different clades.";
RL Infect. Immun. 70:5086-5090(2002).
DR EMBL: AY082389; AAL92494.1; -.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 360
SQ SEQUENCE 360 AA; 39575 MW; 0C09A791547A47EC CRC64;

Query Match 92.7%; Score 468; DB 2; Length 360;
Best Local Similarity 96.2%; Pred. No. 7.6e-24;
Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELDKA-EAELDKKADLPNKVADLEKEISNLEILLG 58
DB 232 LAKKQTELEKLLDLDPEGKTQDELDKAEEAEELDKKADLPNKVADLEKEISNLEILLG 291
QY 59 GADSEDDTAALPNKLTAKKAELEKTKQELDAALNELGPGDDEE 102
DB 292 GADSEDDTAALQNKLTAKKAELEKTKQELDAALNELGPGDDEE 335

RESULT 6
Q9LAX7 PRELIMINARY; PRT; 429 AA.
AC Q9LAX7;
RC STRAIN=AC122;
RX MEDLINE=20448953; PubMed=10992499;
RA DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL: AF071818; AAF27714.1; -.
DR InterPro; IPR000533; Prefoldin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 429
SQ SEQUENCE 429 AA; 47048 MW; BC1D74BBA54DA9D6 CRC64;

Query Match 92.7%; Score 468; DB 2; Length 429;
Best Local Similarity 96.2%; Pred. No. 9e-24;
Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELDKA-EAELDKKADLPNKVADLEKEISNLEILLG 58
DB 254 LAKKQTELEKLLDLDPEGKTQDELDKAEEAEELDKKADLPNKVADLEKEISNLEILLG 313
QY 59 GADSEDDTAALPNKLTAKKAELEKTKQELDAALNELGPGDDEE 102
DB 314 GADSEDDTAALQNKLTAKKAELEKTKQELDAALNELGPGDDEE 357

RESULT 7
Q9LAX9 PRELIMINARY; PRT; 526 AA.
AC Q9LAX9;

```

```

DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF3296;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL: AF071816; AAF27712.1; -.
DR HSP; P04268; IIC2.
DR InterPro; IPR011047; Quin_alc_DH_like.
FT NON_TER 526
SQ SEQUENCE 526 AA; 58106 MW; 5F1F564A2CB678AE CRC64;

Query Match 92.7%; Score 468; DB 2; Length 526;
Best Local Similarity 96.2%; Pred. No. 1.1e-23;
Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELDKA-EAELDKKADLPNKVADLEKEISNLEILLG 58
DB 346 LAKKQTELEKLLDLDPEGKTQDELDKAEEAEELDKKADLPNKVADLEKEISNLEILLG 405
QY 59 GADSEDDTAALPNKLTAKKAELEKTKQELDAALNELGPGDDEE 102
DB 406 GADSEDDTAALQNKLTAKKAELEKTKQELDAALNELGPGDDEE 449

RESULT 8
Q8VQ55 PRELIMINARY; PRT; 608 AA.
AC Q8VQ55;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Pneumococcal surface protein A (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KNIH1156;
RA Lee K.J., Bae S.M., Chung K.S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF460993; AAL67804.1; -.
DR HSP; P06653; 1HCX.
DR Pfam; PF01473; CW binding 1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_2.
FT NON_TER 608
FT NON_TER 608
SQ SEQUENCE 608 AA; 67918 MW; 15F71BD62E297526 CRC64;

Query Match 92.7%; Score 468; DB 2; Length 608;
Best Local Similarity 96.2%; Pred. No. 1.3e-23;
Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELDKA-EAELDKKADLPNKVADLEKEISNLEILLG 58
DB 222 LAKKQTELEKLLDLDPEGKTQDELDKAEEAEELDKKADLPNKVADLEKEISNLEILLG 281
QY 59 GADSEDDTAALPNKLTAKKAELEKTKQELDAALNELGPGDDEE 102

```

```
Db 282 GADSEDDTAALQNKATKKALEKTQKELDAALNELGPDGDEE 325
RESULT 9
Q97T39 PRELIMINARY; PRT; 744 AA.
AC Q97T39;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pneumococcal surface protein A.
GN OrderedLocuNames=SP0117;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916; DOI=10.1126/science.1061217;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S.N., Heidelberg J.F., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Mayhew L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzaple E.K., Khouli H.M., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S.V., Dickinson T.,
RA Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae."
RL Science 293:498-506(2001).
DR EMBL: AE007328; AAK74303.1; -.
DR PIR: F95013; F95013.
DR HSP: P06653; IHXC.
DR TIGR: SP0117; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam: PF01473; CW binding 1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_2.
KW Complete proteome.
SQ SEQUENCE 744 AA; 82764 MW; 20EASE8E7911EFD5 CRC64;

Query Match 92.7%; Score 468; DB 2; Length 744;
Best Local Similarity 96.2%; Pred. No. 1.6e-23;
Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Qy 1 LAKKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPKNKADLEKEINLEILLG 58
Db 346 LAKKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPKNKADLEKEINLEILLG 405

Qy 59 GADSEDDTAALPNKATKKALEKTQKELDAALNELGPDGDEE 102
Db 406 GADSEDDTAALQNKATKKALEKTQKELDAALNELGPDGDEE 449

RESULT 10
Q9L579 PRELIMINARY; PRT; 231 AA.
AC Q9L579;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=20;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF255547; AAF68100.1; -.
DR HSP: P04268; 1IC2.
FT NON_TER 1 231
FT NON_TER 241 241
SQ SEQUENCE 231 AA; 24990 MW; A7731F3A46460186 CRC64;

Query Match 91.7%; Score 463; DB 2; Length 231;
Best Local Similarity 95.2%; Pred. No. 1.1e-23;
Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

Qy 1 LAKKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPKNKADLEKEINLEILLG 58
Db 77 LAKKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPKNKADLEKEINLEILLG 136

Qy 59 GADSEDDTAALPNKATKKALEKTQKELDAALNELGPDGDEE 102
Db 137 GADPEDDTAALQNKATKKALEKTQKELDAALNELGPDGDEE 180

"pneumococcal pspA sequence types of prevalent multiresistant
pneumococcal strains in the United States and of internationally
disseminated clones.";
J. Clin. Microbiol. 38:3663-3669(2000).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=20;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF255548; AAF68101.1; -.
DR HSP: P04268; 1IC2.
FT NON_TER 1 231
FT NON_TER 231 231
SQ SEQUENCE 231 AA; 24990 MW; A7731F3A46460186 CRC64;

Query Match 91.7%; Score 463; DB 2; Length 231;
Best Local Similarity 95.2%; Pred. No. 1.1e-23;
Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

Qy 1 LAKKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPKNKADLEKEINLEILLG 58
Db 76 LAKKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPKNKADLEKEINLEILLG 135

Qy 59 GADSEDDTAALPNKATKKALEKTQKELDAALNELGPDGDEE 102
Db 136 GADPEDDTAALQNKATKKALEKTQKELDAALNELGPDGDEE 179

RESULT 11
Q9L580 PRELIMINARY; PRT; 241 AA.
AC Q9L580;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=121;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=121;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF255547; AAF68100.1; -.
DR HSP: P04268; 1IC2.
FT NON_TER 1 241
FT NON_TER 241 241
SQ SEQUENCE 241 AA; 26038 MW; BB87E1A4C25FA669 CRC64;

Query Match 91.7%; Score 463; DB 2; Length 241;
Best Local Similarity 95.2%; Pred. No. 1.1e-23;
Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

Qy 1 LAKKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPKNKADLEKEINLEILLG 58
Db 77 LAKKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPKNKADLEKEINLEILLG 136

Qy 59 GADSEDDTAALPNKATKKALEKTQKELDAALNELGPDGDEE 102
Db 137 GADPEDDTAALQNKATKKALEKTQKELDAALNELGPDGDEE 180
```

```

RESULT 12
Q9L5B7
ID Q9L5B7 PRELIMINARY; PRT; 249 AA.
AC Q9L5B7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=50;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Packiam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=50;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF253405; AAF67353.1; -.
FT NON_TER 1
FT NON_TER 249
SQ SEQUENCE 249 AA; 27271 MW; B4106707EF108A0B CRC64;

Query Match 91.7%; Score 463; DB 2; Length 249;
Best Local Similarity 95.2%; Pred. No. 1.1e-23;
Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPNKVADLEKEISNLEILG 58
DB 103 LAKKQTELEKLLDLDPEGKTQDELKKEAEEAELDKKADLPNKVADLEKEISNLEILG 162

QY 59 GADSEDDTAALPNKLATTKAELEKTKQKELDAALNELGPDGDEE 102
DB 163 GADPEDDTAALQNKLATTKAELEKTKQKELDAALNELGPDGDEE 206

RESULT 13
Q9LAX8
ID Q9LAX8 PRELIMINARY; PRT; 502 AA.
AC Q9LAX8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG8090;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071817; AAF27713.1; -.
DR HSSP; O15813; 1D7M.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.

FT NON_TER 502
SQ SEQUENCE 502 AA; 55018 MW; 4E73D477CAE79B40 CRC64;

Query Match 90.9%; Score 459; DB 2; Length 502;
Best Local Similarity 94.2%; Pred. No. 4.2e-23;
Matches 98; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPNKVADLEKEISNLEILG 58
DB 347 LAKKQTELEKLLDLDPEGKTQDELKKEAEEAELDKKADLPNKVADLEKEISNLEILG 406

QY 59 GADSEDDTAALPNKLATTKAELEKTKQKELDAALNELGPDGDEE 102
DB 407 GADPEDDTAALQNKLATTKAELEKTKQKELDAALNELGPDGDEE 450

RESULT 14
Q9L585
ID Q9L585 PRELIMINARY; PRT; 249 AA.
AC Q9L585;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=18;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Packiam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=18;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255542; AAF68095.1; -.
DR HSSP; P04268; 1IC2.
FT NON_TER 1
FT NON_TER 249
SQ SEQUENCE 249 AA; 27050 MW; DF4D2ED9265986FA CRC64;

Query Match 90.5%; Score 457; DB 2; Length 249;
Best Local Similarity 94.2%; Pred. No. 2.9e-23;
Matches 98; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPNKVADLEKEISNLEILG 58
DB 68 LAKKQTELEKLLDLDPEGKTQDELKKEAEEAELDKKADLPNKVADLEKEISNLEILG 127

QY 59 GADSEDDTAALPNKLATTKAELEKTKQKELDAALNELGPDGDEE 102
DB 128 GADPEDDTAALQNKLATTKAELEKTKQKELDAALNELGPDGDEE 171

RESULT 15
Q9L590
ID Q9L590 PRELIMINARY; PRT; 256 AA.
AC Q9L590;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:50:39 ; Search time 59.6735 Seconds
(without alignments)
518.502 Million cell updates/sec

Title: US-10-674-755-19

Perfect score: 397

Sequence: 1 LDKEAGEALDKADGLPNK.....TQKELDAALNELGPDGDEE 80

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

- Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	397	100.0	197	7	ABW02598	Abw02598 Ac122c pn
2	397	100.0	8991	6	ABU08487	Abu08487 S. pneumo
3	381.5	96.1	196	2	AAW14564	Aaw14564 Streptoco
4	368	92.7	233	2	AAW14572	Aaw14572 Streptoco
5	368	92.7	233	7	ABW02606	Abw02606 Ef1019c p
6	365	91.9	213	7	ABW02601	Aaw14567 Streptoco
7	363	91.4	213	2	AAW14567	Aaw14567 Streptoco
8	362	91.2	416	8	ADK52498	Adk52498 alpha hel
9	362	91.2	526	6	ADK52497	Adk52497 PspA mole
10	362	91.2	744	6	ABU00449	Abu00449 S. pneumo
11	362	91.2	744	8	ADM2054	Adm2054 S. pneumo
12	358	90.2	745	3	AAW1652	Aay81652 Streptoco
13	358	90.2	641	2	AAW61217	Aaw61217 Streptoco
14	358	90.2	641	5	ABP54636	Abp54636 S. pneumo
15	358	90.2	641	7	ADC45241	Adc45241 S. pneumo
16	238	59.9	211	7	ABW02621	Abw02621 Bg11703c
17	238	59.9	238	2	AAW14587	Aaw14587 Streptoco
18	235	59.2	212	2	AAW14588	Aaw14588 Streptoco
19	235	59.2	212	7	ABW02622	Abw02622 Bg7817c p
20	233	58.7	232	7	ABW02624	Abw02624 Ef5668c p
21	233	58.7	233	2	AAW14590	Aaw14590 Streptoco
22	233	58.7	275	8	ADK52055	Adk52055 S. pneumo
23	233	58.7	369	8	ADK52496	Adk52496 alpha hel
24	233	58.7	458	2	AAW14592	Aaw14592 Streptoco
25	233	58.7	458	7	ABW02626	Abw02626 Ef5668 pn

RESULT 1

ABW02598

ID ABW02598 standard; protein; 197 AA.

XX AC ABW02598;

XX DT 12-FEB-2004 (first entry)

XX DE Ac122c pneumococcal surface protein A (PspA) central region.

XX KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;

XX KW immunological; gene therapy; immunostimulant.

XX OS Unidentified.

XX PN US6592876-B1.

XX PD 15-JUL-2003.

XX PF 15-SEP-1995; 95US-00529055.

XX PR 20-APR-1993; 93US-00048896.

XX PR 06-JUN-1995; 95US-00465746.

XX (UABR-) UAB RES FOUND.

XX PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;

XX WPI; 2003-862841/80.

XX Immunological composition for obtaining expression products used for detecting the presence of Streptococcus pneumoniae or its strain, comprises at least two different full length isolated gene encoding pneumococcal surface protein A.

XX Example 6; SEQ ID NO 44; 121pp; English.

XX The present invention relates to an immunological composition comprising at least 2 different full length isolated genes encoding pneumococcal surface protein A (PspA) from different groups based on restriction fragment polymorphism analysis. The invention is useful for obtaining expression products by recombinant techniques to detect, determine, isolate or diagnose the presence of Streptococcus pneumoniae or its strain. The expression product is useful for preparing antigenic, immunological or vaccine compositions, for eliciting antibodies, an immunological response (other than or additional to antibodies) or a protective response (including antibody or other immunological response by administering compositions to a host). The invention is also useful as

CC vaccines and in gene therapy. The present sequence is Acl22c pneumococcal
 CC surface protein A (PspA) central region. This sequence is used in the
 CC exemplification of the invention

XX Sequence 197 AA;

Query Match 100.0%; Score 397; DB 7; Length 197;
 Best Local Similarity 100.0%; Pred. No. 6.4e-33;
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
 |||||
 Db 46 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 105
 |||||

QY 61 TOKELDAALNELGPDGDEEE 80
 |||||
 Db 106 TOKELDAALNELGPDGDEEE 125
 |||||

RESULT 2
 ID ABU08487
 AC ABU08487 standard; protein; 8991 AA.

XX ABU08487;

XX 24-JUN-2003 (first entry)

XX S. pneumoniae pneumococcal surface protein A (PspA) protein.

XX Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
 KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
 KW antibacterial.

XX Streptococcus pneumoniae.

XX Key Location/Qualifiers

FT Misc-difference 1..8991
 FT /note= "All Xaa residues within this sequence are
 FT unknown"

XX US6500613-B1.

XX 31-DEC-2002.

XX 16-SEP-1996; 96US-00714741.

XX 15-SEP-1995; 95US-00529055.

XX (UYAL-) UNIV ALABAMA.

XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
 PI Hollingshead S, Tart R, Brooks-Walter A;

XX WPI; 2003-361534/34.

XX Isolated PspC amino acid sequence used as polymerase chain reaction or
 PT hybridization probe, comprises pneumococcal surface protein having alpha-
 PT helical, proline rich and repeat regions.

XX Disclosure; Col 145-188; 186pp; English.

XX The present invention relates to the isolation of Streptococcus
 CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
 CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-
 CC like protein having alpha-helical, proline rich and repeat regions. The
 CC PspC and PspA proteins may be used in a vaccine to protect against
 CC pneumococcal infections. The polynucleotide sequences encoding PspC and
 CC PspA may be used for the expression of the proteins, and as PCR primers
 CC or hybridisation probes. The present sequence represents S. pneumoniae
 CC PspA protein

XX Sequence 8991 AA;

Query Match 100.0%; Score 397; DB 6; Length 8991;
 Best Local Similarity 100.0%; Pred. No. 6.8e-31;
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
 |||||
 Db 4108 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 4167
 |||||
 QY 61 TOKELDAALNELGPDGDEEE 80
 |||||
 Db 4168 TOKELDAALNELGPDGDEEE 4187
 |||||

RESULT 3

AAW14564

ID AAW14564 standard; protein; 196 AA.

XX AAW14564;

XX 17-OCT-2003 (revised)

XX 28-OCT-1997 (first entry)

XX Streptococcus pneumoniae PspA central region.

XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
 KW bacteraemia; pneumonia.

XX Streptococcus pneumoniae; strain Acl22.

XX WO9709994-A1.

XX 20-MAR-1997.

XX 16-SEP-1996; 96WO-US014819.

XX 15-SEP-1995; 95US-00529055.

XX (UABR-) UAB RES FOUND.

XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
 PI Hollingshead S, Tart R, Brooks-Walter A;

XX WPI; 1997-202002/18.

XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
 PT in vaccines for protecting animals against S.pneumoniae infection.

XX Example 6; Fig 13; 296pp; English.

XX This sequence shows the central portion, including the C-terminus of the
 CC alpha-helix region and some of the proline-rich region, of pneumococcal
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Acl22.
 CC Comparison of the N-terminal and central regions (AAW14533-57 and
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
 CC be used to divide the strains into several families based on sequence
 CC homologies. PspA polypeptides, or fragments of them, can be used in
 CC vaccines to protect animals against S. pneumoniae infection and hence for
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
 CC region and the immediate 5' tip of the coding sequence are likely to be
 CC the critical sequences for predicting PspA cross-reactions and vaccine
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 196 AA;

Query Match 96.1%; Score 381.5; DB 2; Length 196;
 Best Local Similarity 98.8%; Pred. No. 2.6e-31;
 Matches 79; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
 |||||
 Db 46 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 105
 |||||

QY 61 TOKELDAALNELGPDGDEEE 80
 DB 106 T-KELDAALNELGPDGDEEE 124

RESULT 4
 AAW14572
 ID AAW14572 standard; protein; 233 AA.
 XX
 AC AAW14572;
 XX
 DT 17-OCT-2003 (revised)
 DT 28-OCT-1997 (first entry)
 XX
 DE Streptococcus pneumoniae PspA central region.
 KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
 KW bacteraemia; pneumonia.
 XX
 OS Streptococcus pneumoniae; strain EF3296.
 FH
 FT Key Location/Qualifiers
 FT Misc-difference 129 /note= "unidentified amino acid"
 FT Misc-difference 131 /note= "unidentified amino acid"
 FT
 XX
 PN WO9709994-A1.
 XX
 PD 20-MAR-1997.
 XX
 PF 16-SEP-1996; 96WO-US014819.
 XX
 PR 15-SEP-1995; 95US-00529055.
 XX
 XX (UABR-) UAB RES FOUND.
 XX
 PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
 PI Hollingshead S, Tart R, Brooks-Walter A;
 DR WPI; 1997-202002/18.
 XX
 XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
 PT in vaccines for protecting animals against S.pneumoniae infection.
 XX
 PS Example 6; Fig 13; 296pp; English.
 XX
 CC This sequence shows the central portion, including the C-terminus of the
 CC alpha-helix region and some of the proline-rich region, of pneumococcal
 CC surface protein A (PspA) of Streptococcus pneumoniae strain EF3296.
 CC Comparison of the N-terminal and central regions (AAW14533-57 and
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
 CC be used to divide the strains into several families based on sequence
 CC homologies. PspA polypeptides, or fragments of them, can be used in
 CC vaccines to protect animals against S. pneumoniae infection and hence for
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
 CC region and the immediate 5' tip of the coding sequence are likely to be
 CC the critical sequences for predicting PspA cross-reactions and vaccine
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 233 AA;

Query Match 92.7%; Score 368; DB 2; Length 233;
 Best Local Similarity 93.8%; Pred. No. 7.9e-30;
 Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LDKEAGEALDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
 DB 77 LDKEAGEALDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 136
 QY 61 TOKELDAALNELGPDGDEEE 80
 DB 137 TOKELDAALNELGPDGDEEE 156

DB 137 TOKELDAALNELGPDGDEEE 156

RESULT 5
 ABW02606
 ID ABW02606 standard; protein; 233 AA.
 XX
 AC ABW02606;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 DE Bf1019c pneumococcal surface protein A (PspA) central region.
 XX
 KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
 KW immunological; gene therapy; immunostimulant.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1. .233 /note= "Xaa = Unknown amino acid"
 FT
 XX US6592876-B1.
 XX
 PD 15-JUL-2003.
 XX
 PF 15-SEP-1995; 95US-00529055.
 XX
 PR 20-APR-1993; 93US-00048896.
 PR 06-JUN-1995; 95US-00465746.
 XX
 XX (UABR-) UAB RES FOUND.
 XX
 PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
 XX
 DR WPI; 2003-862841/80.
 XX
 PT Immunological composition for obtaining expression products used for
 PT detecting the presence of Streptococcus pneumoniae or its strain,
 PT comprises at least two different full length isolated gene encoding
 PT pneumococcal surface protein A.
 XX
 PS Example 6; SEQ ID NO 52; 121pp; English.
 XX
 CC The present invention relates to an immunological composition comprising
 CC at least 2 different full length isolated genes encoding pneumococcal
 CC surface protein A (PspAs) from different groups based on restriction
 CC fragment polymorphism analysis. The invention is useful for obtaining
 CC expression products by recombinant techniques to detect, determine,
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its
 CC strain. The expression product is useful for preparing antigenic,
 CC immunological or vaccine compositions, for eliciting antibodies, an
 CC immunological response (other than or additional to antibodies) or a
 CC protective response (including antibody or other immunological response
 CC by administering compositions to a host). The invention is also useful as
 CC vaccines and in gene therapy. The present sequence is Bf1019c
 CC pneumococcal surface protein A (PspA) central region. This sequence is
 CC used in the exemplification of the invention
 XX
 SQ Sequence 233 AA;

Query Match 92.7%; Score 368; DB 7; Length 233;
 Best Local Similarity 93.8%; Pred. No. 7.9e-30;
 Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LDKEAGEALDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
 DB 77 LDKEAGEALDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 136
 QY 61 TOKELDAALNELGPDGDEEE 80
 DB 137 TOKELDAALNELGPDGDEEE 156

```

RESULT 6
ID ABW02601 standard; protein; 213 AA.
XX
AC ABW02601;
XX
DT 12-FEB-2004 (first entry)
XX
DE Bg8090c pneumococcal surface protein A (PspA) central region.
XX
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
XX immunological; gene therapy; immunostimulant.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Misc-difference 2
FT /label= Unknown
XX
PN US6592876-B1.
XX
PD 15-JUL-2003.
XX
PF 15-SEP-1995; 95US-00529055.
XX
PR 20-APR-1993; 93US-00048896.
XX
PR 06-JUN-1995; 95US-00465746.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX WPI; 2003-862841/80.
XX
XX Immunological composition for obtaining expression products used for
XX detecting the presence of Streptococcus pneumoniae or its strain,
XX comprises at least two different full length isolated gene encoding
XX pneumococcal surface protein A.
XX
XX Example 6; SEQ ID NO 47; 121pp; English.
XX
CC The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspAs) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antigenic,
CC immunological or vaccine compositions, for eliciting antibodies, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Bg8090c
CC pneumococcal surface protein A (PspA) central region. This sequence is
CC used in the exemplification of the invention
XX
XX Sequence 213 AA;
XX
Query Match 91.9%; Score 365; DB 7; Length 213;
Best Local Similarity 92.5%; Pred. No. 1.4e-29;
Matches 74; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 LDKEAGEAELDKKADGLPNKVSDELEKEISNLEILLGGADSEDDTAALPNKATKKALEK 60
DB 83 LDKEAAEAELDKKADGLPNKVADLEKEISNLEILLGGADPEDDPTAALPNKATKKAFFEK 142
QY 61 TOKELDAALNELGPDGDEEE 80
DB 143 TPKELDAALNELGPDGDEEE 162
XX
RESULT 7

```

```

AAW14567
ID AAW14567 standard; protein; 213 AA.
XX
AC AAW14567;
XX
DT 17-OCT-2003 (revised)
DT 28-OCT-1997 (first entry)
XX
DE Streptococcus pneumoniae PspA central region.
XX
KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
XX bacteraemia; pneumonia.
XX
OS Streptococcus pneumoniae; strain Bg8090.
XX
FH Key Location/Qualifiers
FT Misc-difference 2
FT /note= "unidentified amino acid"
XX
PN WO9709994-A1.
XX
PD 20-MAR-1997.
XX
PF 16-SEP-1996; 96WO-US014819.
XX
PR 15-SEP-1995; 95US-00529055.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
XX Hollingshead S, Tart R, Brooks-Walter A;
XX WPI; 1997-202002/18.
XX
XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
XX in vaccines for protecting animals against S.pneumoniae infection.
XX
XX Example 6; Fig 13; 296pp; English.
XX
CC This sequence shows the central portion, including the C-terminus of the
CC alpha-helix region and some of the proline-rich region, of pneumococcal
CC surface protein A (PspA) of Streptococcus pneumoniae strain Bg8090.
CC Comparison of the N-terminal and central regions (AAW14533-57 and
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
CC be used to divide the strains into several families based on sequence
CC homologies. PspA polypeptides, or fragments of them, can be used in
CC vaccines to protect animals against S. pneumoniae infection and hence for
CC the prevention of diseases such as otitis media, meningitis, bacteraemia
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
CC region and the immediate 5' tip of the coding sequence are likely to be
CC the critical sequences for predicting PspA cross-reactions and vaccine
CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 213 AA;
XX
Query Match 91.4%; Score 363; DB 2; Length 213;
Best Local Similarity 92.5%; Pred. No. 2.3e-29;
Matches 74; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 LDKEAGEAELDKKADGLPNKVSDELEKEISNLEILLGGADSEDDTAALPNKATKKALEK 60
DB 83 LDKEAAEAELDKKADGLPNKVADLEKEISNLEILLGGADPEDDPTAALPNKATKKAFFEK 142
QY 61 TOKELDAALNELGPDGDEEE 80
DB 143 TPKELDAALNELGPDGDEEE 162
XX
RESULT 8
ID ADK52498 standard; protein; 416 AA.
XX
AC ADK52498;

```


XX 20-MAY-2004 (first entry)
 XX DE alpha helical region PspA molecule from the EF3296 strain.
 XX KW Streptococcus pneumoniae infection; pneumococcal surface protein A; PspA;
 XX KW hemolytic anemia disease; leukemia; lymphoma; sickle cell anemia;
 XX KW Hodgkin's disease.
 XX OS Streptococcus pneumoniae.
 XX PN WO2004016231-A2.
 XX PD 26-FEB-2004.
 XX PF 17-FEB-2003; 2003WO-US008199.
 XX PR 15-MAR-2002; 2002US-0365351P.
 XX PA (UABR-) UAB RES FOUND.
 XX PI Briles DE;
 XX DR WPI; 2004-192068/18.
 XX PT Treating Streptococcus pneumoniae infection in a subject lacking a
 PT functional spleen comprises administering an antibody that recognizes
 PT pneumococcal surface protein A (PspA) or its binding portion.
 XX PS Claim 17; SEQ ID NO 4; 41pp; English.
 XX CC The present invention relates to treating Streptococcus pneumoniae
 CC infection in a subject lacking a functional spleen comprises
 CC administering an antibody that recognizes pneumococcal surface protein A
 CC (PspA) or its binding portion. The method is useful for treating or
 CC preventing Streptococcus pneumoniae infection in a subject lacking a
 CC functional spleen. The disease-associated injury is especially due to
 CC hemolytic anemia disease, leukemia or lymphoma, especially sickle cell
 CC anemia or Hodgkin's disease. The present sequence represents the alpha
 CC helical region PspA molecule from the EF3296 strain of Streptococcus
 CC pneumoniae.
 XX SQ Sequence 416 AA;
 Query Match 91.2%; Score 362; DB 8; Length 416;
 Best Local Similarity 93.8%; Pred. No. 6.6e-29;
 Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 LDKEAGEALDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
 DB 265 LDKEAEAEALDKKADGLQNKVADLEKEISNLEILLGGADSEDDTAALQNKLTATKKALEK 324
 QY 61 TOKELDAALNELGPDGDEE 80
 DB 325 TOKELDAALNELGPDGDEE 344
 RESULT 9
 ADK52497
 ID ADK52497 standard; protein; 526 AA.
 XX AC ADK52497;
 XX DT 20-MAY-2004 (first entry)
 XX DE PspA molecule from the EF3296 strain of Streptococcus pneumoniae.
 XX KW Streptococcus pneumoniae infection; pneumococcal surface protein A; PspA;
 XX KW hemolytic anemia disease; leukemia; lymphoma; sickle cell anemia;
 XX KW Hodgkin's disease.
 XX OS Streptococcus pneumoniae.
 XX PN WO2004016231-A2.
 XX PD 03-OCT-2002.
 XX PF 27-MAR-2002; 2002WO-IB002163.
 XX PR 27-MAR-2001; 2001GB-00007658.
 XX PA (CHIR-) CHIRON SPA.
 XX PA (GENO-) INST GENOMIC RES.

PN WO2004016231-A2.
 XX PD 26-FEB-2004.
 XX PF 17-FEB-2003; 2003WO-US008199.
 XX PR 15-MAR-2002; 2002US-0365351P.
 XX PA (UABR-) UAB RES FOUND.
 XX PI Briles DE;
 XX DR WPI; 2004-192068/18.
 XX PT Treating Streptococcus pneumoniae infection in a subject lacking a
 PT functional spleen comprises administering an antibody that recognizes
 PT pneumococcal surface protein A (PspA) or its binding portion.
 XX PS Claim 17; SEQ ID NO 3; 41pp; English.
 XX CC The present invention relates to treating Streptococcus pneumoniae
 CC infection in a subject lacking a functional spleen comprises
 CC administering an antibody that recognizes pneumococcal surface protein A
 CC (PspA) or its binding portion. The method is useful for treating or
 CC preventing Streptococcus pneumoniae infection in a subject lacking a
 CC functional spleen. The disease-associated injury is especially due to
 CC hemolytic anemia disease, leukemia or lymphoma, especially sickle cell
 CC anemia or Hodgkin's disease. The present sequence represents PspA
 CC molecule from the EF3296 strain of Streptococcus pneumoniae.
 XX SQ Sequence 526 AA;
 Query Match 91.2%; Score 362; DB 8; Length 526;
 Best Local Similarity 93.8%; Pred. No. 8.9e-29;
 Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 LDKEAGEALDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
 DB 370 LDKEAEAEALDKKADGLQNKVADLEKEISNLEILLGGADSEDDTAALQNKLTATKKALEK 429
 QY 61 TOKELDAALNELGPDGDEE 80
 DB 430 TOKELDAALNELGPDGDEE 449
 RESULT 10
 ABU00449
 ID ABU00449 standard; protein; 744 AA.
 XX AC ABU00449;
 XX DT 23-OCT-2003 (revised)
 DT 11-FEB-2003 (first entry)
 XX DE S. pneumoniae type 4 strain protein from coding region #16.
 XX KW Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;
 KW antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;
 KW gene therapy; vaccine.
 XX OS Streptococcus pneumoniae; type 4 strain.
 XX PN WO200277021-A2.
 XX PD 03-OCT-2002.
 XX PF 27-MAR-2002; 2002WO-IB002163.
 XX PR 27-MAR-2001; 2001GB-00007658.
 XX PA (CHIR-) CHIRON SPA.
 XX PA (GENO-) INST GENOMIC RES.

```
PI Masignani V, Tettelin H, Fraser C;
XX WPI; 2003-040579/03.
DR N-PSDB; ABX05728.
DR
XX
XX
XX New proteins and nucleic acid molecules from Streptococcus pneumoniae,
PT useful as medicaments for treating or preventing a disease or infection
PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or
PT ear infection.
XX
XX Claim 1; SEQ ID NO 32; 56pp; English.
PS
XX
XX The invention relates to a protein comprising or having at least 50%
CC identity to any of the 2469 amino acid sequences, identified in the
CC specification (available on a computer readable format), or its fragment,
CC expressed from 2469 of 2489 identified DNA coding regions from the
CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
CC ADS56454. Also included are an antibody which binds one of the proteins,
CC treating a patient by administering the protein, DNA or antibody (in a
CC composition), a kit comprising first and second primers, which are the
CC nucleic acid cited above or fragments between nucleotides 8-100 of a
CC sequence not defined in the specification, for amplifying a target
CC sequence contained within a Streptococcus nucleic acid sequence, where
CC the first primer is substantially complementary to the target sequence
CC and the second primer is substantially complementary to the complement of
CC the target sequence, and where the parts of the primers having
CC substantial complementarity define the termini of the target sequence to
CC be amplified, assay comprising contacting a test compound with the
CC protein, and determining whether the test compound binds to the protein
CC and a Streptococcus pneumoniae bacterium, where one or more genes
CC encoding the proteins has been rendered inactive. The proteins, nucleic
CC acid molecules, antibody and compositions are useful as medicaments for
CC treating or preventing a disease or infection due to streptococcus
CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis
CC media or ear infection. They are also useful in developing vaccines,
CC diagnostics and antibiotics. The methods are useful for identifying
CC immunodominant proteins. The present sequence is one of the 2469 proteins
CC expressed by the identified coding regions from the genomic sequence.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to
XX standardise OS field)
XX
SQ Sequence 744 AA;

Query Match          91.2%; Score 362; DB 6; Length 744;
Best Local Similarity 93.8%; Pred. No. 1.4e-28;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 370 LDKEAEAEALDKKADGLQNKVADLEKEISNLEILLGGADSEDDTAALQNKLTAKKAELEK 429

QY 61 TQKELDAALNELGPDGDEEE 80
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 430 TQKELDAALNELGPDGDEEE 449

RESULT 12
AAY81652
ID AAY81652 standard; protein; 745 AA.
XX
AC AAY81652;
XX
DT 24-MAY-2000 (first entry)
XX
DE Streptococcus pneumoniae protein sequence ID301.
XX
XX Streptococcus pneumoniae; vaccine; screening; protein antigen;
XX antibacterial; antiinflammatory; meningitis; infection; diagnosis;
XX pneumococcal disease.
XX
OS Streptococcus pneumoniae.
XX
PN WO200006737-A2.
XX
PD 10-FEB-2000.
XX
PF 27-JUL-1999; 99WO-GB002451.
XX
PR 27-JUL-1998; 98GB-00016337.
PR 19-MAR-1999; 99US-0125164P.
XX
PA (MICR-) MICROBIAL TECHNICS LTD.
XX
PI Gilbert CFG, Hansbro PM;
XX
```

PI Masignani V, Tettelin H, Fraser C;
XX WPI; 2003-040579/03.
DR N-PSDB; ABX05728.
DR
XX
XX
XX New proteins and nucleic acid molecules from Streptococcus pneumoniae,
PT useful as medicaments for treating or preventing a disease or infection
PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or
PT ear infection.
XX
XX Claim 1; SEQ ID NO 32; 56pp; English.
PS
XX
XX The invention relates to a protein comprising or having at least 50%
CC identity to any of the 2469 amino acid sequences, identified in the
CC specification (available on a computer readable format), or its fragment,
CC expressed from 2469 of 2489 identified DNA coding regions from the
CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
CC ADS56454. Also included are an antibody which binds one of the proteins,
CC treating a patient by administering the protein, DNA or antibody (in a
CC composition), a kit comprising first and second primers, which are the
CC nucleic acid cited above or fragments between nucleotides 8-100 of a
CC sequence not defined in the specification, for amplifying a target
CC sequence contained within a Streptococcus nucleic acid sequence, where
CC the first primer is substantially complementary to the target sequence
CC and the second primer is substantially complementary to the complement of
CC the target sequence, and where the parts of the primers having
CC substantial complementarity define the termini of the target sequence to
CC be amplified, assay comprising contacting a test compound with the
CC protein, and determining whether the test compound binds to the protein
CC and a Streptococcus pneumoniae bacterium, where one or more genes
CC encoding the proteins has been rendered inactive. The proteins, nucleic
CC acid molecules, antibody and compositions are useful as medicaments for
CC treating or preventing a disease or infection due to streptococcus
CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis
CC media or ear infection. They are also useful in developing vaccines,
CC diagnostics and antibiotics. The methods are useful for identifying
CC immunodominant proteins. The present sequence is one of the 2469 proteins
CC expressed by the identified coding regions from the genomic sequence.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to
XX standardise OS field)
XX
SQ Sequence 744 AA;

Query Match 91.2%; Score 362; DB 6; Length 744;
Best Local Similarity 93.8%; Pred. No. 1.4e-28;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 370 LDKEAEAEALDKKADGLQNKVADLEKEISNLEILLGGADSEDDTAALQNKLTAKKAELEK 429

QY 61 TQKELDAALNELGPDGDEEE 80
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 430 TQKELDAALNELGPDGDEEE 449

RESULT 11
ADM92054
ID ADM92054 standard; protein; 744 AA.
XX
AC ADM92054;
XX
DT 03-JUN-2004 (first entry)
XX
XX S pneumoniae antigenic protein sequence SeqID251.
DE
DE antibacterial; gene therapy; Streptococcus pneumoniae infection;
XX antigenic.
XX
XX Streptococcus pneumoniae.
OS
PI

DR WPI; 2000-195300/17.
 XX New Streptococcal protein, useful as a vaccine, for diagnosis of
 PT pneumococcal diseases and for screening agents capable of antagonizing or
 PT inhibiting expression of the protein.
 XX Claim 2; Page 95; 108pp; English.
 XX
 CC AAY81501 to AAY81679 represent specifically claimed protein sequences
 CC isolated from Streptococcus pneumoniae. AAA05407 to AAA05590 represent
 CC specifically claimed nucleotide sequences isolated from S. pneumoniae.
 CC The sequences have antibacterial and antiinflammatory properties. The
 CC protein sequences, and fragments of them, are useful as immunogens and/or
 CC antigens. The nucleotide sequences can be used in vaccines and in
 CC diagnostic assays. The proteins and nucleotides can be useful for the
 CC detection and diagnosis of S. pneumoniae. The protein sequences are also
 CC useful for screening an agent capable of antagonising, inhibiting or
 CC interfering with the function or expression of the proteins in which the
 CC agent is useful for treatment or prophylaxis of S. pneumoniae infection
 CC and meningitis. AAA05591 to AAA05614 represent primers used in the
 CC exemplification of the present invention
 XX
 SQ Sequence 745 AA;
 Query Match 91.2%; Score 362; DB 3; Length 745;
 Best Local Similarity 93.8%; Pred. No. 1.4e-28;
 Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 LDKAGEAELDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
 DB 370 LDKAGEAELDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 429
 QY 61 TQKELDAALNELGPDGDEEE 80
 DB 430 TQKELDAALNELGPDGDEEE 449
 RESULT 13
 AAW61217
 ID AAW61217 standard; protein; 641 AA.
 XX
 AC AAW61217;
 XX
 DT 02-OCT-1998 (first entry)
 DE
 DE Streptococcus pneumoniae SP0092 protein.
 XX
 XX Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;
 KW Streptococcus pneumoniae; otitis media; meningitis.
 XX
 OS Streptococcus pneumoniae.
 XX
 XX
 PH Key Location/Qualifiers
 FT Misc-difference 306
 FT /label= unknown
 FT /note= "encoded by NCT"
 XX
 XX WO9818930-A2.
 XX
 PD 07-MAY-1998.
 XX
 XX 30-OCT-1997; 97WO-US019422.
 XX
 PR 31-OCT-1996; 96US-0029960P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Kunsch CA, Choi GH, Johnson LS, Hromockyj A;
 PI
 PI WPI; 1998-272224/24.
 DR N-PSDB; AAV27403.
 DR
 XX Nucleic acid encoding antigenic peptide(s) from Streptococcus pneumoniae

PT - or their epitope-containing fragments, useful in protective or
 therapeutic vaccines, and for diagnosis.
 XX Claim 11; Page 82; 118pp; English.
 XX
 CC The present sequence represents a protein from Streptococcus pneumoniae.
 CC The nucleic acid sequence encoding the Streptococcus pneumoniae protein
 CC can be useful in vaccines for inducing protective antibodies against
 CC Streptococcus pneumoniae, for treatment or prevention of infection e.g.
 CC pneumonia, otitis media or meningitis. Probes based on the nucleic acid
 CC are used to detect Streptococcus infection (by usual hybridisation or
 CC amplification methods), also for isolating Streptococcus genes or their
 CC allelic variants. The protein can be used similarly to detect specific
 CC antibodies in standard immunoassays, especially for diagnosing or
 CC monitoring infections. Antibodies which bind the protein are used to
 CC detect corresponding antigens, to purify the protein and for passive
 CC immunisation (optionally coupled to a toxin). Vaccines are administered,
 CC e.g. by injection, orally or through the skin, typically at 0.01-1000
 CC (especially 10-300) mu g/ml per dose
 XX
 SQ Sequence 641 AA;
 Query Match 90.2%; Score 358; DB 2; Length 641;
 Best Local Similarity 92.5%; Pred. No. 2.9e-28;
 Matches 74; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 1 LDKAGEAELDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
 DB 267 LDKAGEAELDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 326
 QY 61 TQKELDAALNELGPDGDEEE 80
 DB 327 TQKELDAALNELGPDGDEEE 346
 RESULT 14
 ABP54636
 ID ABP54636 standard; protein; 641 AA.
 XX
 AC ABP54636;
 XX
 DT 04-SEP-2002 (first entry)
 DE
 DE S. pneumoniae SP092 protein sequence SEQ ID NO:160.
 XX
 KW Streptococcus pneumoniae; epitope; vaccine; antigenic protein;
 KW antibacterial; Streptococcal infection; detection.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN US2002061545-A1.
 XX
 PD 23-MAY-2002.
 XX
 PF 22-JAN-2001; 2001US-00765272.
 XX
 PR 30-OCT-1997; 97US-00961083.
 XX
 PA (CHOI/) CHOI G H.
 PA (KUNS/) KUNSCH C A.
 PA (BARA/) BARASH S C.
 PA (DILL/) DILLON P J.
 PA (DOUG/) DOUGHERTY B.
 PA (FANN/) FANNON M R.
 PA (ROSE/) ROSEN C A.
 XX
 XX Choi GH, Kunsch CA, Barash SC, Dillon PJ, Dougherty B, Fannon MR;
 PI Rosen CA;
 PI WPI; 2002-479261/51.
 DR N-PSDB; ABQ84871.
 DR
 XX New Streptococcus pneumoniae antigens, useful for detecting Streptococcus

PT and for preventing or attenuating disease caused by Streptococcus
infection.

XX
PS Claim 11; Page 43; 70pp; English.

XX ABQ84792 to ABQ84904 represents nucleic acids which encode the
CC Streptococcus pneumoniae antigens given in ABP54557 to ABP54669. The S.
CC pneumoniae antigens have antibacterial activity and can be used in
CC vaccines. The S. pneumoniae antigens can also be used to prevent or
CC attenuate a Streptococcal infection in an animal. The polynucleotides
CC encoding the S. pneumoniae antigens can be used to detect Streptococcus
CC nucleic acids. ABQ84905 to ABQ85130 represent primers used in the cloning
CC of S. pneumoniae ORFs (open reading frames) which are used in an example
CC from the present invention

XX
SQ Sequence 641 AA;

Query Match 90.2%; Score 358; DB 5; Length 641;
Best Local Similarity 92.5%; Pred. No. 2.9e-28;
Matches 74; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LDKKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
|||||
Db 267 LDKKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 326

QY 61 TOKELDAALNELGPDGDEE 80
|||||
Db 327 TOKELDAALNELGPDGDEE 346

RESULT 15
ADC45241
ID ADC45241 standard; protein; 641 AA.

AC ADC45241;

XX 18-DEC-2003 (first entry)

XX S. pneumoniae antigenic protein SP092.

XX Antigen; bacterial infection; vaccine; pneumonia; antibacterial.

XX Streptococcus pneumoniae.

XX US6573082-B1.

XX 03-JUN-2003.

XX 28-MAR-2000; 2000US-00536784.

XX 31-OCT-1996; 96US-0029960P.

XX 30-OCT-1997; 97US-00961083.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Choi GH, Kunsch CA, Barash SC, Dillon PJ, Dougherty B, Fannon MR;
PI Roen CA;

XX WPI; 2003-764574/72.

XX N-PSDB; ADC45240.

XX Novel polynucleotide encoding Streptococcus pneumoniae polypeptides
PT useful for producing vaccines for prevention or attenuation of infection
PT by Streptococcus pneumoniae.

XX Example 1; SEQ ID NO 160; 58pp; English.

XX The invention relates to an isolated polynucleotide consisting of a
CC Streptococcus pneumoniae nucleic acid (appearing as ADC45122 and encoding
CC SP028) one of 113 disclosed nucleic acids encoding 113 S. pneumoniae
CC antigens. Also included are making a recombinant vector by inserting the
CC nucleic acid into a vector, an isolated polynucleotide consisting of at
CC least 50 or 100 contiguous nucleotides of the SP028 nucleic acid, and a

CC recombinant host cell comprising the SP028 polynucleotide. The nucleic
CC acids are useful as DNA vaccine against Streptococcus pneumoniae
CC infection (e.g. pneumonia). Nucleic acids derived from the S. pneumoniae
CC antigen nucleic acids are useful as probes for use in diagnostic methods
CC for detecting S. pneumoniae gene expression. The present sequence
XX represents an S. pneumoniae antigenic protein.

SQ Sequence 641 AA;

Query Match 90.2%; Score 358; DB 7; Length 641;
Best Local Similarity 92.5%; Pred. No. 2.9e-28;
Matches 74; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LDKKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
|||||
Db 267 LDKKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 326

QY 61 TOKELDAALNELGPDGDEE 80
|||||
Db 327 TOKELDAALNELGPDGDEE 346

Search completed: June 21, 2005, 10:10:17
Job time : 59.6735 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:55:14 ; Search time 14.9388 Seconds
(without alignments)
399.760 Million cell updates/sec

Title: US-10-674-755-19
Perfect score: 397
Sequence: 1 LDKAGEALDKKADGLPNK.....TQKELDAALNELGPDGDEEE 80

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B COMB.pep:*
3: /cgn2_6/prodata/1/iaa/6A COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6B COMB.pep:*
5: /cgn2_6/prodata/1/iaa/PCTUS COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	397	100.0	80	2	US-08-710-749-18
2	397	100.0	80	4	US-09-147-875A-19
3	397	100.0	197	4	US-08-529-055-44
4	397	100.0	8991	4	US-08-714-741-32
5	378	95.2	104	2	US-08-710-749-19
6	378	95.2	104	4	US-09-147-875A-20
7	375	94.5	104	4	US-09-147-875A-21
8	369.5	93.1	102	2	US-08-710-749-21
9	369.5	93.1	102	4	US-08-710-749-21
10	369	92.9	104	2	US-08-147-875A-18
11	368	92.7	233	4	US-08-710-749-20
12	365	91.9	213	4	US-08-529-055-52
13	358	90.2	641	3	US-08-529-055-47
14	358	90.2	641	4	US-08-529-055-47
15	238	59.9	108	2	US-08-710-749-22
16	238	59.9	108	2	US-08-710-749-23
17	238	59.9	108	2	US-08-710-749-26
18	238	59.9	108	4	US-09-147-875A-23
19	238	59.9	211	4	US-08-529-055-67
20	235	59.2	108	4	US-09-147-875A-24
21	235	59.2	212	4	US-08-529-055-68
22	233	58.7	108	2	US-08-710-749-24
23	233	58.7	108	4	US-09-147-875A-25
24	233	58.7	232	4	US-08-529-055-70
25	233	58.7	458	4	US-08-529-055-73
26	225	56.7	106	4	US-08-147-875A-22
27	186	46.9	108	2	US-08-710-749-25

28	186	46.9	108	4	US-09-147-875A-26	Sequence 26, Appl
29	186	46.9	185	4	US-08-529-055-69	Sequence 69, Appl
30	149.5	37.7	119	2	US-08-710-749-27	Sequence 27, Appl
31	149.5	37.7	119	4	US-09-147-875A-27	Sequence 27, Appl
32	149.5	37.7	215	4	US-08-529-055-43	Sequence 43, Appl
33	109.5	27.6	204	4	US-08-529-055-51	Sequence 51, Appl
34	106.5	26.8	195	4	US-08-529-055-71	Sequence 71, Appl
35	106	26.7	623	4	US-08-714-741-47	Sequence 47, Appl
36	105.5	26.6	99	2	US-08-710-749-10	Sequence 10, Appl
37	105.5	26.6	99	2	US-08-710-749-11	Sequence 11, Appl
38	105.5	26.6	99	4	US-09-147-875A-11	Sequence 61, Appl
39	105.5	26.6	198	4	US-08-529-055-61	Sequence 61, Appl
40	105.5	26.6	288	3	US-08-312-949-4	Sequence 4, Appl
41	105.5	26.6	288	3	US-08-446-201-4	Sequence 4, Appl
42	105.5	26.6	289	1	US-08-072-070-4	Sequence 4, Appl
43	105.5	26.6	289	1	US-08-469-434-4	Sequence 4, Appl
44	105.5	26.6	289	1	US-08-214-222-4	Sequence 4, Appl
45	105.5	26.6	289	2	US-08-467-852A-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-710-749-18
; Sequence 18, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-18

Query Match 100.0%; Score 397; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 8.3e-37;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LDKAGEALDKKADGLPNKVSLEKEISNLIILGGADSEDDTAALPNKLTAKKAELEK 60
DB 1 LDKAGEALDKKADGLPNKVSLEKEISNLIILGGADSEDDTAALPNKLTAKKAELEK 60


```

Query Match      \ 100.0%; Score 397; DB 4; Length 8991;
Best Local Similarity 100.0%; Pred. No. 3.9e-34;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LDKKAEAEALDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
DB 4108 LDKKAEAEALDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
DB 4168 TQKELDAALNELGPDGDEEE 80
DB 4168 TQKELDAALNELGPDGDEEE 4187

RESULT 5
US-08-710-749-19
; Sequence 19, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESS: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-19

Query Match      95.2%; Score 378; DB 2; Length 104;
Best Local Similarity 96.2%; Pred. No. 1.5e-34;
Matches 77; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LDKKAEAEALDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
DB 25 LDKKAEAEALDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 84
DB 61 TQKELDAALNELGPDGDEEE 80
DB 85 TQKELDAALNELGPDGDEEE 104

RESULT 6
US-09-147-875A-20
; Sequence 20, Application US/09147875A
; Patent No. 6638516

```

```
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 102 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-710-749-21

Query Match 93.1%; Score 369.5; DB 2; Length 102;
Best Local Similarity 96.2%; Pred. No. 1.3e-33;
Matches 77; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 LDKEAGEAELDKKADGPNKVSDEKEISNLEILGGADSEDDTAALPNKLTAKAELEK 60
DB 24 LDKEA-EAELDKKADLPNKVADLEKEISNLEILGGADSEDDTAALPNKLTAKAELEK 82
QY 61 TQKELDAALNELGPDGDEEE 80
DB 83 TQKELDAALNELGPDGDEEE 102

RESULT 9
US-09-147-875A-18
; Sequence 18, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-147-875A-18

Query Match 93.1%; Score 369.5; DB 4; Length 102;
Best Local Similarity 96.2%; Pred. No. 1.3e-33;
Matches 77; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 LDKEAGEAELDKKADGPNKVSDEKEISNLEILGGADSEDDTAALPNKLTAKAELEK 60
DB 24 LDKEA-EAELDKKADLPNKVADLEKEISNLEILGGADSEDDTAALPNKLTAKAELEK 82
QY 61 TQKELDAALNELGPDGDEEE 80
DB 83 TQKELDAALNELGPDGDEEE 102

RESULT 10
```

```
US-08-710-749-20
; Sequence 20, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-710-749-20

Query Match 92.9%; Score 369; DB 2; Length 104;
Best Local Similarity 93.8%; Pred. No. 1.5e-33;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LDKEAGEAELDKKADGPNKVSDEKEISNLEILGGADSEDDTAALPNKLTAKAELEK 60
DB 25 LDKEA-EAELDKKADLPNKVADLEKEISNLEILGGADSEDDTAALPNKLTAKAELEK 84
QY 61 TQKELDAALNELGPDGDEEE 80
DB 85 TQKELDAALNELGPDGDEEE 104

RESULT 11
US-08-529-055-52
; Sequence 52, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
```


CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/529,055
FILING DATE: 15-SEP-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454312-2400
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 233 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-529-055-52

[illegible]

RESULT 12
US-08-529-055-47
Sequence 47, Application US/08529055
Patent No. 6592876
GENERAL INFORMATION:
APPLICANT: Briles, David E.
APPLICANT: McDaniel, Larry S.
APPLICANT: Swiatlo, Edwin
APPLICANT: Yother, Janet
APPLICANT: Brooks-Walter, Alexis
TITLE OF INVENTION: Pneumococcal Genes, Portions
TITLE OF INVENTION: Thereof, Expression Products
TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
TITLE OF INVENTION: Portions and Products
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford, P.C.
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/529,055
FILING DATE: 15-SEP-1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454312-2400
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 213 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-0A-529-055-47

```

Query Match      91.9%; Score 365; DB 4; Length 213;
Best Local Similarity 92.5%; Pred. No. 1e-32;
Matches 74; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LDKEAGEAELDKKADGLPNKVSLEKISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
    |||||
Db 83 LDKEAAEAELDKKADGLPNKVADLEKISNLEILLGGADPEDDTAALPNKLTATKKALEK 142
    |||||

Qy 61 TQKELDAALNELGPDGDEEE 80
    |||||
Db 143 TPKELEDAALNELGPDGDEEE 162
    |||||

```

RESULT 13
US-08-961-083-160
Sequence 160, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

```

ZIF: 40830
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 160:
SEQUENCE CHARACTERISTICS:
LENGTH: 641 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-961-083-160

```

Query Match 90.2%; Score 358; DB 3; Length 641;

RESULT 2

```
US-10-299-636-59
; Sequence 59, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR FILING DATE: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-59

Query Match          100.0%; Score 397; DB 15; Length 197;
Best Local Similarity 100.0%; Pred. No. 2e-31;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 46 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 105

Qy 61 TQKELDAALNELGPDGDEE 80
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 106 TQKELDAALNELGPDGDEE 125

RESULT 3
US-10-674-755-20
; Sequence 20, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-20

Query Match          95.2%; Score 378; DB 15; Length 104;
Best Local Similarity 96.2%; Pred. No. 7e-30;
Matches 77; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 84

Qy 61 TQKELDAALNELGPDGDEE 80
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85 TQKELDAALNELGPDGDEE 104

US-10-674-755-21
; Sequence 21, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-21

Query Match          94.5%; Score 375; DB 15; Length 104;
Best Local Similarity 95.0%; Pred. No. 1.4e-29;
Matches 76; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 84

Qy 61 TQKELDAALNELGPDGDEE 80
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85 TQKELDAALNELGPDGDEE 104

RESULT 5
US-10-674-755-18
; Sequence 18, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-18

Query Match          93.1%; Score 369.5; DB 15; Length 102;
Best Local Similarity 96.2%; Pred. No. 4.7e-29;
Matches 77; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 24 LDKEA-EAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 82

Qy 61 TQKELDAALNELGPDGDEE 80
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 83 TQKELDAALNELGPDGDEE 102

RESULT 6
US-10-299-636-67
; Sequence 67, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
```

```
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 67
; LENGTH: 233
; TYPE: PR
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (129)
; OTHER INFORMATION: Xaa at position 129 is unknown
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (131)
; OTHER INFORMATION: Xaa at position 131 is unknown
US-10-299-636-67

Query Match          92.7%; Score 368; DB 15; Length 233;
Best Local Similarity 93.8%; Pred. No. 1.9e-28;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      1 LDKEAGEAELDKKADGLPNKVSDLEKEISNLEILLGGADSEDDTAALPNKLATKKAELEK 60
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      77 LDKEAGEAELDKKADGLPNKVADLEKEISNLEILLGGADSEDDTAALPNKLAKXCAELEK 136
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY      61 TOKELDAALNELGPDGDEEE 80
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      137 TQKELDAAPNELGPDGDEEE 156
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
US-10-299-636-62
; Sequence 62, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 213
; TYPE: PR
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
```


QY 57 ELEKTOKELDAALNELGPDGDEEE 80
Db 109 ELEKTPKELDAALNELGPDGDEEE 132

RESULT 13
US-10-674-755-24
; Sequence 24, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-24

Query Match 59.2%; Score 235; DB 15; Length 108;
Best Local Similarity 63.1%; Pred. No. 1.1e-15;
Matches 53; Conservative 11; Mismatches 16; Indels 4; Gaps 2;

QY 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADS---ED-DTAALPNKLATKKA 56
Db 25 LDKEAAEAEELNKVEALPNQVAEELEELSKELEDNLKDAETNNVEDYIKGLEEBAIATKQA 84

QY 57 ELEKTOKELDAALNELGPDGDEEE 80
Db 85 ELEKTPKELDAALNELGPDGDEEE 108

RESULT 14
US-10-299-636-83
; Sequence 83, Application US/10299636
; Publication No. US2004007847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: Briles, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 83
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-83

Query Match 59.2%; Score 235; DB 15; Length 212;
Best Local Similarity 63.1%; Pred. No. 2.5e-15;
Matches 53; Conservative 11; Mismatches 16; Indels 4; Gaps 2;

QY 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADS---ED-DTAALPNKLATKKA 56

Db 52 LDKEAAEAEELNKVEALPNQVAEELEELSKELEDNLKDAETNNVEDYIKGLEEBAIATKQA 111
QY 57 ELEKTOKELDAALNELGPDGDEEE 80
Db 112 ELEKTPKELDAALNELGPDGDEEE 135

RESULT 15
US-10-674-755-25
; Sequence 25, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-25

Query Match 58.7%; Score 233; DB 15; Length 108;
Best Local Similarity 63.1%; Pred. No. 1.7e-15;
Matches 53; Conservative 11; Mismatches 16; Indels 4; Gaps 2;

QY 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADS---ED-DTAALPNKLATKKA 56
Db 25 LDKEAAEAEELNKVEALPNQVAEELEELSKELEDNLKDAETNNVEDYIKGLEEBAIATKKA 84

QY 57 ELEKTOKELDAALNELGPDGDEEE 80
Db 85 ELEKTPKELDAALNELGPDGDEEE 108

Search completed: June 21, 2005, 11:18:37
Job time : 51.102 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:53:39 ; Search time 8 Seconds
(without alignments)
962.168 Million cell updates/sec

Title: US-10-674-755-19
Perfect score: 397
Sequence: 1 LDKEAGEAELDKADGLPNK.....TQKELDAALNELGPDGDEEE 80

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	362	91.2	744	2 F95013	pneumococcal surfa
2	105.5	26.6	619	2 A97887	surface protein ps
3	105.5	26.6	619	2 A41971	surface protein ps
4	89.5	22.5	284	2 S19691	tropomyosin alpha,
5	88.5	22.3	284	2 JC6199	alpha-tropomyosin
6	87	21.9	415	2 S35760	fcrA protein precu
7	86.5	21.8	284	2 A45488	body-wall muscle t
8	85.5	21.5	284	2 JC6198	alpha-tropomyosin
9	85	21.4	405	2 A33939	Fc gamma (IgG) rec
10	84.5	21.3	281	2 A34787	tropomyosin 1 alph
11	84.5	21.3	284	1 TMRBA	tropomyosin alpha
12	84.5	21.3	284	2 A39816	tropomyosin 2, fib
13	84.5	21.3	284	2 B27407	tropomyosin alpha
14	84.5	21.3	284	2 A27674	tropomyosin 3, fib
15	84.5	21.3	284	2 A25825	tropomyosin alpha
16	84.5	21.3	284	2 JC2551	tropomyosin alpha
17	84.5	21.3	284	2 A60597	tropomyosin 2, fib
18	84.5	21.3	284	2 B39816	tropomyosin 3, fib
19	84.5	21.3	2139	2 T18296	myosin heavy chain
20	84	21.2	1509	1 A27224	myosin heavy chain
21	83	20.9	388	2 A46173	Mrp4 protein - Str
22	83	20.9	516	2 B84709	hypothetical prote
23	83	20.9	1837	2 T41023	probable nuclear p
24	82.5	20.8	280	2 A22165	tropomyosin alpha
25	82.5	20.8	284	2 S24972	tropomyosin alpha,
26	82	20.7	936	2 S39083	myosin heavy chain
27	81.5	20.5	308	2 T08796	tropomyosin - huma
28	81.5	20.5	587	2 JC1419	Fc gamma (IgG) rec
29	81.5	20.5	1051	2 T18302	apsB protein - Eme

ALIGNMENTS

RESULT 1

F95013
pneumococcal surface protein A [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C:Accession: F95013
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapf, I.
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: F95013
A>Status: Preliminary
A:Molecule type: DNA
A:Cross-references: 1-744 <KUR>
A:Cross-references: UNIPROT:Q97T39; GB:AE005672; PIDN:AAK74303.1; PID:gl4971584; GSPDB:G
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP0117

Query Match 91.2%; Score 362; DB 2; Length 744;
Best Local Similarity 93.8%; Pred. No. 2e-22;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LDKEAGEAELDKADGLPNKVSLEKEISNLEILIGGADSEDDTAALPNKLTAKAELEK 60
DB 370 LDKEAGEAELDKADGLPNKVSLEKEISNLEILIGGADSEDDTAALPNKLTAKAELEK 429

QY 61 TQKELDAALNELGPDGDEEE 80
DB 430 TQKELDAALNELGPDGDEEE 449

RESULT 2

surface protein pspA precursor [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: A97887
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: A97887
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-619 <KUR>

A:Residues: 1-415 <POD>
A:Cross-references: UNIPROT:Q54859; EMBL:X69324; NID:g311759; PIDN:CAA49165.1; PID:g3117
R:Haanes, E.J.; Heath, D.G.; Cleary, P.P.
J. Bacteriol. 174, 4967-4976, 1992
A:Title: Architecture of the vir regulons of group A streptococci parallels opacity fact
A:Reference number: A42711; MUID:92332431; PMID:1385809
A:Accession: A42711
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 343-415 <HAA>
A:Cross-references: GB:M86806; NID:g153630; PIDN:AAA26887.1; PID:g153631
A:Experimental source: strain CS101, OP+
A:Note: sequence extracted from NCBI backbone (NCBIN:108942, NCBIP:108945)
C:Superfamily: M5 protein

Query Match 21.9%; Score 87; DB 2; Length 415;
Best Local Similarity 33.3%; Pred. No. 4.9;
Matches 28; Conservative 12; Mismatches 28; Indels 16; Gaps 3;

QY 3 KGAEGAELDKKADGLPNK---VSDLEKEISNLEILLGG-----ASEDDTAAL 47
DB 125 KAAAEAAQKALDALNNKKNQISDLTNEAQLKEAIEGYVQTIONASRBIAAQOELAAV 184
QY 48 PNKLATKKALEKTQKELDAALNE 71
DB 185 KSQLEAKNAEIDL-KQQDASKTE 207

RESULT 7
A45488
body-wall muscle tropomyosin - sea squirt (Ciona intestinalis)
C:Species: Ciona intestinalis
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A45488
R:Meedel, T.H.; Hastings, K.E.
J. Biol. Chem. 268, 6735-6764, 1993
A:Title: Striated muscle-type tropomyosin in a chordate smooth muscle, ascidian body-wall
A:Reference number: A45488
A:Accession: A45488
A:Molecule type: nucleic acid
A:Residues: 1-284 <NEE>
A:Cross-references: UNIPROT:Q07068; GB:X64105; NID:g297457; PIDN:CAA45469.1; PID:g297458
A:Note: sequence extracted from NCBI backbone (NCBIN:128007, NCBIP:128008)
C:Superfamily: tropomyosin

Query Match 21.8%; Score 86.5; DB 2; Length 284;
Best Local Similarity 29.7%; Pred. No. 3.7;
Matches 30; Conservative 16; Mismatches 26; Indels 29; Gaps 4;

QY 1 LDKE-----AGEAELDK---KADGLPNKVSLEKE-----ISNLE 32
DB 13 LDKENAIADRAEQARTDKSAEDKATGLBEELQGLKRLKATEDELTSQEKLRATLENLE 72
QY 33 -ILGGADSEDDTAALPNKLATKKALEKTQKELDAALNEL 72
DB 73 NAEKKAADAQEAQEVASLNRRITLVEELDRAQERLTISLSKL 113

RESULT 8
JC6198
alpha-tropomyosin C-2 - axolotl
C:Species: Ambystoma mexicanum (axolotl)
C:Date: 11-Apr-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004
C:Accession: JC6198
R:Luque, E.A.; Spinner, B.J.; Dube, S.; Dube, D.K.; Lemanski, L.F.
Gene 185, 175-180, 1997
A:Title: Differential expression of a novel isoform of alpha-tropomyosin in cardiac and
A:Reference number: JC6198; MUID:97208870; PMID:9055812
A:Contents: Heart
A:Accession: JC6198
A:Molecule type: mRNA
A:Residues: 1-284 <LUQ>

A:Cross-references: UNIPROT:P87348; GB:U33449; NID:gi1871355; PIDN:AAC60091.1; PID:gi187135
C:Comment: This protein is a actin-binding protein.
C:Genetics:
A:Gene: AtmC-2
C:Superfamily: tropomyosin
C:Keywords: actin binding

Query Match 21.5%; Score 85.5; DB 2; Length 284;
Best Local Similarity 29.1%; Pred. No. 4.4;
Matches 30; Conservative 12; Mismatches 28; Indels 33; Gaps 3;

QY 1 LDKE-----AGEAELDKKADGLPNKVSLEKEISNLEILL----- 35
DB 13 IDKENAMRAEQEADKK--GAEDSKQLESEIVQLEKQLRISEDRDRVLDLHKSEES 70
QY 36 -----GGADSEDDTAALPNKLATKKALEKTQKELDAALNEL 72
DB 71 LITADEKAAKGSDDAASLNRRITQLVEEELDRAQERLATALQKL 113

RESULT 9
A33939
Fc gamma (IgG) receptor II precursor - Streptococcus sp. (fragment)
C:Species: Streptococcus sp.
C:Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 26-Aug-1999
C:Accession: A33939
R:Heath, D.G.; Cleary, P.P.
Proc. Natl. Acad. Sci. U.S.A. 86, 4741-4745, 1989
A:Title: Fc-receptor and M-protein genes of group A streptococci are products of gene dup
A:Reference number: A33939; MUID:89282846; PMID:2660147
A:Accession: A33939
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-405 <HEA>
A:Cross-references: GB:M22532; NID:g153628; PIDN:AAB95296.1; PID:g552003
C:Superfamily: M5 protein
C:Keywords: immunoglobulin receptor

Query Match 21.4%; Score 85; DB 2; Length 405;
Best Local Similarity 34.2%; Pred. No. 7;
Matches 25; Conservative 12; Mismatches 28; Indels 8; Gaps 2;

QY 3 KGAEGAELDKKADGLPNK---VSDLEKEISNLEILLGG-----ASEDDTAALPNKLATK 54
DB 135 KAAAEAAQKALDALNNKKNQISDLTNEAQLKEAIEGYVQTIONASRBIAAQOELAAA 194
QY 55 KAELEKTQKELDA 67
DB 195 KSQLEAKNAEIEA 207

RESULT 10
A34787
tropomyosin 1 alpha, brain - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 13-Jul-1990 #sequence_revision 13-Jul-1990 #text_change 09-Jul-2004
C:Accession: A34787
R:Lees-Miller, J.P.; Goodwin, L.O.; Helfman, D.M.
Mol. Cell. Biol. 10, 1729-1742, 1990
A:Title: Three novel brain tropomyosin isoforms are expressed from the rat alpha-tropomy
A:Reference number: A34787; MUID:90205854; PMID:2320008
A:Accession: A34787
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-281 <LEE>
A:Cross-references: UNIPROT:P04692; GB:M34135
C:Superfamily: tropomyosin
C:Keywords: alternative splicing; coiled coil

Query Match 21.3%; Score 84.5; DB 2; Length 281;
Best Local Similarity 28.2%; Pred. No. 5.3;
Matches 29; Conservative 15; Mismatches 26; Indels 33; Gaps 3;

Search completed: June 21, 2005, 10:12:04
Job time : 8 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:54:24 ; Search time 49.551 Seconds
(without alignments)
826.751 Million cell updates/sec

Title: US-10-674-755-19

Perfect score: 397

Sequence: 1 LDKAGEAELDKKADGLPNK.....TQKELDAALNELGPDGDEEE 80

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	362	91.2	228	2 Q9L5B8	Q9L5B8 streptococc
2	362	91.2	235	2 Q9L582	Q9L582 streptococc
3	362	91.2	249	2 Q9L5D4	Q9L5D4 streptococc
4	362	91.2	252	2 Q9L583	Q9L583 streptococc
5	362	91.2	360	2 Q8KQK3	Q8KQK3 streptococc
6	362	91.2	429	2 Q9LAX7	Q9LAX7 streptococc
7	362	91.2	526	2 Q9LAX9	Q9LAX9 streptococc
8	362	91.2	608	2 Q8VQ55	Q8VQ55 streptococc
9	362	91.2	744	2 Q97T39	Q97T39 streptococc
10	359	90.4	231	2 Q9L579	Q9L579 streptococc
11	359	90.4	241	2 Q9L580	Q9L580 streptococc
12	358	90.2	242	2 Q9L562	Q9L562 streptococc
13	357	89.9	249	2 Q9L5B7	Q9L5B7 streptococc
14	355	89.4	502	2 Q9LAX8	Q9LAX8 streptococc
15	353	88.9	209	2 Q9L593	Q9L593 streptococc
16	353	88.9	249	2 Q9L585	Q9L585 streptococc
17	353	88.9	256	2 Q9L590	Q9L590 streptococc
18	342	86.1	222	2 Q9L584	Q9L584 streptococc
19	237	59.7	246	2 Q9L5B4	Q9L5B4 streptococc
20	233	58.7	479	2 Q9LAX2	Q9LAX2 streptococc
21	233	58.7	481	2 Q9LAX5	Q9LAX5 streptococc
22	233	58.7	653	2 Q34097	Q34097 streptococc
23	232	58.4	107	2 Q8KQK2	Q8KQK2 streptococc
24	221	55.7	213	2 Q8GNS7	Q8GNS7 streptococc
25	208	52.4	480	2 Q9LAX3	Q9LAX3 streptococc
26	159.5	40.2	211	2 Q8GNT0	Q8GNT0 streptococc
27	159.5	40.2	227	2 Q9KGS0	Q9KGS0 streptococc
28	159.5	40.2	256	2 Q9L595	Q9L595 streptococc
29	159.5	40.2	257	2 Q9L594	Q9L594 streptococc
30	159.5	40.2	461	2 Q9LAX6	Q9LAX6 streptococc
31	109.5	27.6	417	2 Q9LAY3	Q9LAY3 streptococc

ALIGNMENTS

RESULT 1

Q9L5B8 PRELIMINARY; PRT; 228 AA.
AC Q9L5B8;
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=60;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Cherardi G., Packlam R.R., Hollingshead S.K.;
RT "pneumococcal pspA sequence types of prevalent multiresistant
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=60;
RX Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF253404; AAF67352.1; -;
DR InterPro; IPR009053; Prefoldin.
FT NON_TER 1
FT NON_TER 228
SQ SEQUENCE 228 AA; 24430 MW; E6EAA953EC54EA0F CRC64;
Query Match 91.2%; Score 362; DB 2; Length 228;
Best Local Similarity 93.8%; Pred.No.8.3e-22;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

RESULT 2

Q9L582 PRELIMINARY; PRT; 235 AA.
AC Q9L582;
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE PspA (Fragment).
QY 1 LDKAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKATKKALEK 60
Db 70 LDKAEEAELDKKADGLQNKVADLEKEISNLEILLGGADSEDDTAALQNKATKKALEK 129
QY 61 TQKELDAALNELGPDGDEEE 80
Db 130 TQKELDAALNELGPDGDEEE 149

```
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3;
RA Beall B.W.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=3;
RA Beall B.W.;
RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255545; AAF68098.1; -.
FT NON_TER 1 1
FT NON_TER 235 235
SQ SEQUENCE 235 AA; 25424 MW; BFFB48C52CA8380 CRC64;

Query Match 91.2%; Score 362; DB 2; Length 235;
Best Local Similarity 93.8%; Pred. No. 8.6e-22;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LDKEAGEAELDKADGLPNKVSDEKEISNLEILLGGADEDDTAALPNKLTATKKALEK 60
Db 80 LDKEAGEAELDKADELQNKVADLEKEISNLEILLGGADEDDTAALQNKLTATKKALEK 139
QY 61 TQKELDAALNELGPDGDEEE 80
Db 140 TQKELDAALNELGPDGDEEE 159

RESULT 3
Q9L5D4 PRELIMINARY; PRT; 249 AA.
AC Q9L5D4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP195;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SP195;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF252286; AAF69499.1; -.
FT NON_TER 1 1
FT NON_TER 249 249
SQ SEQUENCE 249 AA; 26621 MW; F8EA39225CF8D43F CRC64;

Query Match 91.2%; Score 362; DB 2; Length 249;
Best Local Similarity 93.8%; Pred. No. 9.1e-22;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3;
RA Beall B.W.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=3;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF252286; AAF69499.1; -.
FT NON_TER 1 1
FT NON_TER 235 235
SQ SEQUENCE 235 AA; 25424 MW; BFFB48C52CA8380 CRC64;

Query Match 91.2%; Score 362; DB 2; Length 235;
Best Local Similarity 93.8%; Pred. No. 8.6e-22;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LDKEAGEAELDKADGLPNKVSDEKEISNLEILLGGADEDDTAALPNKLTATKKALEK 60
Db 80 LDKEAGEAELDKADELQNKVADLEKEISNLEILLGGADEDDTAALQNKLTATKKALEK 139
QY 61 TQKELDAALNELGPDGDEEE 80
Db 140 TQKELDAALNELGPDGDEEE 159

RESULT 4
Q9L583 PRELIMINARY; PRT; 252 AA.
AC Q9L583;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=127;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=127;
RA Beall B.W.;
RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255544; AAF68097.1; -.
FT NON_TER 1 1
FT NON_TER 252 252
SQ SEQUENCE 252 AA; 27260 MW; 82DE13741F369CA2 CRC64;

Query Match 91.2%; Score 362; DB 2; Length 252;
Best Local Similarity 93.8%; Pred. No. 9.2e-22;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LDKEAGEAELDKADGLPNKVSDEKEISNLEILLGGADEDDTAALPNKLTATKKALEK 60
Db 97 LDKEAGEAELDKADELQNKVADLEKEISNLEILLGGADEDDTAALQNKLTATKKALEK 156
QY 61 TQKELDAALNELGPDGDEEE 80
Db 157 TQKELDAALNELGPDGDEEE 176

RESULT 5
Q8KQK3 PRELIMINARY; PRT; 360 AA.
AC Q8KQK3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Pneumococcal surface protein A (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=259/98;
RX MEDLINE=22170754; PubMed=12183557;
```



```
RX DOI=10.1128/IAI.70.9.5086-5090.2002;
RA Miyaji E.N., Ferreira D.M., Lopes A.P.Y., Brandileone M.C.C.,
RA Dias W.O., Leite L.C.C.;
RT "Analysis of serum cross-reactivity and cross-protection elicited by
RT immunization with DNA vaccines against Streptococcus pneumoniae
RT expressing PspA fragments from different clades.";
RL Infect. Immun. 70:5086-5090(2002).
DR EMBL: AY082389; AAL92494.1; -.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 360
SQ SEQUENCE 360 AA; 39575 MW; 0C09A791547A47EC CRC64;

Query Match 91.2%; Score 362; DB 2; Length 360;
Best Local Similarity 93.8%; Pred. No. 1.3e-21;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LDKEAGEAELDKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
DB 256 LDKEAEAEELDKADELQNKVADLEKEISNLEILLGGADSEDDTAALQNKLTATKKALEK 315
QY 61 TQKELDAALNELGPDGDEE 80
DB 316 TQKELDAALNELGPDGDEE 335

RESULT 6
Q9LAX7 PRELIMINARY; PRT; 429 AA.
AC Q9LAX7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AC122;
RX MEDLINE=20448953; PubMed=10992499;
RA DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL: AF071816; AAF27712.1; -.
DR HSSP; P04268; IIC2.
DR InterPro; IPR011047; Quin_abc_DH_like.
FT NON_TER 526
FT NON_TER 526
SQ SEQUENCE 526 AA; 58106 MW; 5F1F564A2CB678AE CRC64;

Query Match 91.2%; Score 362; DB 2; Length 526;
Best Local Similarity 93.8%; Pred. No. 2e-21;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LDKEAGEAELDKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
DB 370 LDKEAEAEELDKADELQNKVADLEKEISNLEILLGGADSEDDTAALQNKLTATKKALEK 429
QY 61 TQKELDAALNELGPDGDEE 80
DB 430 TQKELDAALNELGPDGDEE 449

RESULT 8
Q8VQ55 PRELIMINARY; PRT; 608 AA.
AC Q8VQ55;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pneumococcal surface protein A (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KNIH1156;
RA Lee K.J., Bae S.W., Chung K.S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF460993; AAL67804.1; -.
DR HSSP; P06653; IHCX.
DR Pfam; PF01473; CW binding 1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_2.
FT NON_TER 1
FT NON_TER 608
FT NON_TER 608
SQ SEQUENCE 608 AA; 67918 MW; 15F71BD62E297526 CRC64;

Query Match 91.2%; Score 362; DB 2; Length 608;
Best Local Similarity 93.8%; Pred. No. 2.3e-21;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LDKEAGEAELDKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
DB 246 LDKEAEAEELDKADELQNKVADLEKEISNLEILLGGADSEDDTAALQNKLTATKKALEK 305
QY 61 TQKELDAALNELGPDGDEE 80
DB 430 TQKELDAALNELGPDGDEE 449

RESULT 9
Q9LAX9 PRELIMINARY; PRT; 526 AA.
AC Q9LAX9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KNIH1156;
RA Lee K.J., Bae S.W., Chung K.S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF460993; AAL67804.1; -.
DR HSSP; P06653; IHCX.
DR Pfam; PF01473; CW binding 1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_2.
FT NON_TER 1
FT NON_TER 526
FT NON_TER 526
SQ SEQUENCE 526 AA; 58106 MW; 5F1F564A2CB678AE CRC64;

Query Match 91.2%; Score 362; DB 2; Length 526;
Best Local Similarity 93.8%; Pred. No. 2e-21;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LDKEAGEAELDKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
DB 246 LDKEAEAEELDKADELQNKVADLEKEISNLEILLGGADSEDDTAALQNKLTATKKALEK 305
QY 61 TQKELDAALNELGPDGDEE 80
DB 430 TQKELDAALNELGPDGDEE 449
```

Db 306 TQKELDAALNELGPDGDEE 325

RESULT 9

Q97T39 PRELIMINARY; PRT; 744 AA.

ID Q97T39 SEQUENCE FROM N.A.

AC Q97T39; STRAIN=20;

DT 01-OCT-2001 (TrEMBLrel. 18, Created)

DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Pneumococcal surface protein A.

GN OrderedLocusNames=SP0117;

OS Streptococcus pneumoniae.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=1313;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC BAA-334 / TIGR4;

RX MEDLINE=21357209; PubMed=11463916; DOI=10.1126/science.1061217;

RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D., Peterson S.N., Heidelberg J.F., DeBoy R.T., Haft D.H., Dodson R.J., Durkin A.S., Gwinn M.L., Kolonay J.F., Nelson W.C., Peterson J.D., Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radue D., Holtzapple E.K., Khouri H.M., Wolf A.M., Utterback T.R., Hansen C.L., McDonald L.A., Feldblyum T.V., Angiuoli S.V., Dickinson T., Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C., Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.; "Complete genome sequence of a virulent isolate of Streptococcus pneumoniae";

RT Science 293:498-506 (2001).

RL EMBL; AE007328; AAK74303.1; -.

DR PIR; F95013; F95013.

DR HSP; P06653; LHGX.

DR TIGR; SP0117; -.

DR InterPro; IPR002479; CW_binding.

DR InterPro; IPR002345; Lipocalin.

DR Pfam; PF01473; CW_binding_1; 10.

DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_2.

KW Complete proteome.

SQ SEQUENCE 744 AA; 82764 MW; 20EA5E8E7911EFDS CRC64;

Query Match 91.2%; Score 362; DB 2; Length 744;

Best Local Similarity 93.8%; Pred. No. 2.9e-21;

Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LDKEAGEALDKADGLPNKYSVLEKISNLEILLGGADSDDDTAALPNKLTAKAELEK 60

DB 370 LDKEAGEALDKADGLPNKYSVLEKISNLEILLGGADSDDDTAALPNKLTAKAELEK 429

QY 61 TQKELDAALNELGPDGDEE 80

DB 430 TQKELDAALNELGPDGDEE 449

RESULT 10

Q9L579 PRELIMINARY; PRT; 231 AA.

ID Q9L579 SEQUENCE FROM N.A.

AC Q9L579; STRAIN=121;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE PspA (Fragment).

GN Name=pspA;

OS Streptococcus pneumoniae.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=1313;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=121;

RA Beall B.W.;

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF255547; AAF68100.1; -.

DR HSP; P04268; 1IC2.

FT NON_TER 1 241

FT NON_TER 241 241

SQ SEQUENCE 231 AA; 24990 MW; A7731F3A46460186 CRC64;

Query Match 90.4%; Score 359; DB 2; Length 231;

Best Local Similarity 92.5%; Pred. No. 1.5e-21;

Matches 74; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LDKEAGEALDKADGLPNKYSVLEKISNLEILLGGADSDDDTAALPNKLTAKAELEK 60

DB 101 LDKEAGEALDKADGLPNKYSVLEKISNLEILLGGADSDDDTAALPNKLTAKAELEK 160

QY 61 TQKELDAALNELGPDGDEE 80

DB 161 TQKELDAALNELGPDGDEE 180

RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;

```

RESULT 12
Q9L562
ID Q9L562 PRELIMINARY; PRT; 242 AA.
AC Q9L562;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE PepA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=69;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=69;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255908; AAF70098.1;
FT NON_TER 1
FT NON_TER 242
SQ SEQUENCE 242 AA; 25843 MW; 707BA930797D2C82 CRC64;

Query Match 90.2%; Score 358; DB 2; Length 242;
Best Local Similarity 91.2%; Pred. No. 1.9e-21;
Matches 73; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKKEAGEALDKKADGLPNKVSDELEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
Db 94 LKKEAAEALDKKADGLPNKVSDELEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 153
QY 61 TOKELDAALNELGPDGDEEE 80
Db 154 TOKELDAALNELGPDGDEEE 173

RESULT 13
Q9L5B7
ID Q9L5B7 PRELIMINARY; PRT; 249 AA.
AC Q9L5B7;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE PepA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=50;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=50;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

Query Match 89.4%; Score 355; DB 2; Length 502;
Best Local Similarity 91.2%; Pred. No. 7e-21;
Matches 73; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKKEAGEALDKKADGLPNKVSDELEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
Db 371 LKKEAAEALDKKADGLPNKVSDELEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 430
QY 61 TOKELDAALNELGPDGDEEE 80
Db 431 TOKELDAALNELGPDGDEEE 450

RESULT 14
Q9LAX8
ID Q9LAX8 PRELIMINARY; PRT; 502 AA.
AC Q9LAX8;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE PepA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG8090;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of pspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae."
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071817; AAF27713.1;
DR HSSP; O15813; 1D7M.
DR InterPro; IPR011047; Quin_ald_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PRO0194; TROPOMYOSIN.
FT NON_TER 502
SQ SEQUENCE 502 AA; 55018 MW; 4E73D477CAE79B40 CRC64;

Query Match 89.4%; Score 355; DB 2; Length 502;
Best Local Similarity 91.2%; Pred. No. 7e-21;
Matches 73; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKKEAGEALDKKADGLPNKVSDELEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
Db 371 LKKEAAEALDKKADGLPNKVSDELEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 430
QY 61 TOKELDAALNELGPDGDEEE 80
Db 431 TOKELDAALNELGPDGDEEE 450

RESULT 15
Q9L593
ID Q9L593 PRELIMINARY; PRT; 209 AA.
AC Q9L593;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE PepA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.

```

```

OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=115;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Packlam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=115;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF254256; AAF68091.1; -.
FT NON_TER 1
FT NON_TER 209
SQ SEQUENCE 209 AA; 22628 MW; 06FF588F7C3BD5B7 CRC64;

Query Match      88.9%; Score 353; DB 2; Length 209;
Best Local Similarity 91.2%; Pred.No. 4.1e-21;
Matches 73; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LDKEAGEAELDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
Db 43 LDKEAAEAEELDKKADGLQNKVADLEKEISNLEILLGGADPEDDTAALQNKLTATKKALEK 102

Qy 61 TQKELDAALNELGPDGDEEE 80
Db 103 TQKELDAALNELGPDGDEEE 122

```

Search completed: June 21, 2005, 10:22:14
Job time : 50.551 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:50:39 ; Search time 77.5755 Seconds
(without alignments)
518.502 Million cell updates/sec

Title: US-10-674-755-20
Perfect score: 514
Sequence: 1 LAKQTELEKLLDSLDPEGK.....TQKELDAALNBLPGDDEE 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	498	96.9	416	8	Adk52498 alpha hel
2	498	96.9	526	8	Adk52497 PepA mole
3	498	96.9	744	6	Abu00449 S. pneumo
4	498	96.9	744	8	Adm92054 S. pneumon
5	498	96.9	745	3	Aay81652 Streptoco
6	494	96.1	641	2	Aaw61217 Streptoco
7	494	96.1	641	5	Abp54636 S. pneumo
8	494	96.1	641	7	Adc45241 S. pneumo
9	490	95.3	213	7	Abw02601 Bg8090c p
10	490	95.3	8991	6	Abu08487 S. pneumo
11	488	94.9	213	2	Aaw14567 Streptoco
12	486	94.6	197	7	Abw02598 Ac122c pn
13	465	90.5	233	7	Abw02606 Ef1019c p
14	462.5	90.0	196	2	Aaw14564 Streptoco
15	458	89.1	233	2	Aaw14572 Streptoco
16	319	62.1	211	7	Abw02621 Bg11703c
17	319	62.1	238	2	Aaw14587 Streptoco
18	313	60.9	232	7	Abw02624 Ef5668c p
19	313	60.9	275	8	Ado52055 S. pneumo
20	313	60.9	369	8	Adk52496 alpha hel
21	313	60.9	458	2	Aaw14592 Streptoco
22	313	60.9	458	7	Abw02626 Ef5668 pn
23	313	60.9	653	8	Adk52495 PepA mole
24	313	60.9	653	8	Ado52080 S. pneumo
25	304	59.1	212	2	Aaw14588 Streptoco

26	304	59.1	212	7	ABW02622	Abw02622 Bg7817c p
27	302.5	58.9	233	2	AAW14590	Aaw14590 Streptoco
28	277	53.9	459	8	ADO15316	Ado15316 S_pneumon
29	256	49.8	185	7	ABW02623	Abw02623 Bg7561c p
30	242.5	47.2	184	2	AAW14589	Aaw14589 Streptoco
31	198	38.5	487	8	ADR04321	Adr04321 Streptoco
32	198	38.5	489	8	ADO52088	Ado52088 Streptoco
33	198	38.5	524	8	ADO52082	Ado52082 E. coli B
34	198	38.5	627	8	ADO52129	Ado52129 E. coli B
35	193.5	37.6	119	2	AAW46291	Aaw46291 Pneumono
36	193.5	37.6	215	7	AAW14563	Aaw14563 Streptoco
37	193.5	37.6	215	7	ABW02597	Abw02597 Atcc6303c
38	185.5	36.1	290	8	ADO52119	Ado52119 pYA3637 b
39	185.5	36.1	298	8	ADO52127	Ado52127 pYA3637 b
40	180	35.0	230	8	ADO52086	Ado52086 S. pneumo
41	180	35.0	230	8	ADR04319	Adr04319 Streptoco
42	129.5	25.2	550	8	ADK48356	Adk48356 Streptoco
43	129.5	25.2	550	8	ADR95223	Adr95223 Novel S.
44	128.5	25.0	289	2	AAW62276	Aaw62276 Streptoco
45	128.5	25.0	289	2	AAW41840	Aaw41840 Streptoco

ALIGNMENTS

RESULT 1

ADK52498
ID ADK52498 standard; protein; 416 AA.

XX ADK52498;

XX 20-MAY-2004 (first entry)

XX alpha helical region PspA molecule from the EF3296 strain.

XX Streptococcus pneumoniae infection; pneumococcal surface protein A; PspA;
KW hemolytic anemia disease; leukemia; lymphoma; sickle cell anemia;
KW Hodgkin's disease.

XX Streptococcus pneumoniae.

XX WO2004016231-A2.

XX 26-FEB-2004.

XX 17-FEB-2003; 2003WO-US008199.

XX 15-MAR-2002; 2002US-0365351P.

XX (UABR-) UAB RES FOUND.

XX Briles DE;

XX WPI; 2004-192068/18.

XX Treating Streptococcus pneumoniae infection in a subject lacking a functional spleen comprises administering an antibody that recognizes pneumococcal surface protein A (PspA) or its binding portion.

XX Claim 17; SEQ ID NO 4; 41pp; English.

XX The present invention relates to treating Streptococcus pneumoniae infection in a subject lacking a functional spleen comprises administering an antibody that recognizes pneumococcal surface protein A (PspA) or its binding portion. The method is useful for treating or preventing Streptococcus pneumoniae infection in a subject lacking a functional spleen. The disease-associated injury is especially due to hemolytic anemia disease, leukemia or lymphoma, especially sickle cell anemia or Hodgkin's disease. The present sequence represents the alpha helical region PspA molecule from the EF3296 strain of Streptococcus pneumoniae.

XX Sequence 416 AA;

SQ

```
Query Match          96.9%; Score 498; DB 8; Length 416;
Best Local Similarity 98.1%; Pred. No. 1.9e-35;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDLSLDPGKGTQDELQKEAEAEALDKKADLPKNKQVADLEKEISNLEILLG 60
   |||||||
Db 241 LAKKQTELEKLLDLSLDPGKGTQDELQKEAEAEALDKKADLPKNKQVADLEKEISNLEILLG 300
   |||||||

QY 61 GADSEDDTAALPNKLTAKKAELEKTQKELDAAALNELGPDGDEEE 104
   |||||||
Db 301 GADSEDDTAALQNLKATKKALEKTQKELDAAALNELGPDGDEEE 344
   |||||||

RESULT 2
ADK52497
ID ADK52497 standard; protein; 526 AA.
XX AC ADK52497;
XX DT 20-MAY-2004 (first entry)
XX DE PepA molecule from the EF3296 strain of Streptococcus pneumoniae.
XX KW Streptococcus pneumoniae infection; pneumococcal surface protein A; PspA;
XX KW hemolytic anemia disease; leukemia; lymphoma; sickle cell anemia;
XX KW Hodgkin's disease.
XX OS Streptococcus pneumoniae.
XX PN WO2004016231-A2.
XX PD 26-FEB-2004.
XX PF 17-FEB-2003; 2003WO-US008199.
XX PR 15-MAR-2002; 2002US-0365351P.
XX PA (UABR-) UAB RES FOUND.
XX PI Briles DE;
XX DR WPI; 2004-192068/18.
XX PT Treating Streptococcus pneumoniae infection in a subject lacking a
XX PT functional spleen comprises administering an antibody that recognizes
XX PT pneumococcal surface protein A (PspA) or its binding portion.
XX PS Claim 17; SEQ ID NO 3; 41pp; English.
XX CC The present invention relates to treating Streptococcus pneumoniae
XX CC infection in a subject lacking a functional spleen comprises
XX CC administering an antibody that recognizes pneumococcal surface protein A
XX CC (PspA) or its binding portion. The method is useful for treating or
XX CC preventing Streptococcus pneumoniae infection in a subject lacking a
XX CC functional spleen. The disease-associated injury is especially due to
XX CC hemolytic anemia disease, leukemia or lymphoma, especially sickle cell
XX CC anemia or Hodgkin's disease. The present sequence represents PspA
XX CC molecule from the EF3296 strain of Streptococcus pneumoniae.
XX SQ Sequence 526 AA;

Query Match          96.9%; Score 498; DB 8; Length 526;
Best Local Similarity 98.1%; Pred. No. 2.4e-35;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDLSLDPGKGTQDELQKEAEAEALDKKADLPKNKQVADLEKEISNLEILLG 60
   |||||||
Db 346 LAKKQTELEKLLDLSLDPGKGTQDELQKEAEAEALDKKADLPKNKQVADLEKEISNLEILLG 405
   |||||||

QY 61 GADSEDDTAALPNKLTAKKAELEKTQKELDAAALNELGPDGDEEE 104
   |||||||
Db 406 GADSEDDTAALQNLKATKKALEKTQKELDAAALNELGPDGDEEE 449
   |||||||
```

```
RESULT 3
ABU00449
ID ABU00449 standard; protein; 744 AA.
XX AC ABU00449;
XX DT 23-OCT-2003 (revised)
XX DT 11-FEB-2003 (first entry)
XX DE S. pneumoniae type 4 strain protein from coding region #16.
XX KW Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;
XX KW antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;
XX KW gene therapy; vaccine.
XX OS Streptococcus pneumoniae; type 4 strain.
XX PN WO200277021-A2.
XX PD 03-OCT-2002.
XX PF 27-MAR-2002; 2002WO-IB002163.
XX PR 27-MAR-2001; 2001GB-00007658.
XX PA (CHIR-) CHIRON SPA.
XX PA (GENO-) INST GENOMIC RES.
XX PI Masignani V, Tettelin H, Fraser C;
XX DR WPI; 2003-040579/03.
XX DR N-PSDB; ABX05728.
XX PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,
XX PT useful as medicaments for treating or preventing a disease or infection
XX PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or
XX PT ear infection.
XX PS Claim 1; SEQ ID NO 32; 56pp; English.
XX CC The invention relates to a protein comprising or having at least 50%
XX CC identity to any of the 2469 amino acid sequences, identified in the
XX CC specification (available on a computer readable format), or its fragment,
XX CC expressed from 2469 of 2489 identified DNA coding regions from the
XX CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
XX CC ABS56454. Also included are an antibody which binds one of the proteins,
XX CC treating a patient by administering the protein, DNA or antibody (in a
XX CC composition), a kit comprising first and second primers, which are the
XX CC nucleic acid cited above or fragments between nucleotides 8-100 of a
XX CC sequence not defined in the specification, for amplifying a target
XX CC sequence contained within a Streptococcus nucleic acid sequence, where
XX CC the first primer is substantially complementary to the target sequence, and
XX CC the second primer is substantially complementary to the complement of
XX CC the target sequence, and where the parts of the primers having
XX CC substantial complementarity define the termini of the target sequence to
XX CC be amplified, assay comprising contacting a test compound with the
XX CC protein, and determining whether the test compound binds to the protein
XX CC and a Streptococcus pneumoniae bacterium, where one or more genes
XX CC encoding the proteins has been rendered inactive. The proteins, nucleic
XX CC acid molecules, antibody and compositions are useful as medicaments for
XX CC treating or preventing a disease or infection due to streptococcus
XX CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis
XX CC media or ear infection. They are also useful in developing vaccines,
XX CC diagnostics and antibiotics. The methods are useful for identifying
XX CC immunodominant proteins. The present sequence is one of the 2469 proteins
XX CC expressed by the identified coding regions from the genomic sequence.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-Oct-2003 to
XX CC standardise OS field)
```

```
SQ Sequence 744 AA;
Query Match 96.9%; Score 498; DB 6; Length 744;
Best Local Similarity 98.1%; Pred. No. 3.6e-35;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDLDPEGKTQDELKAEAEELDKKADELPNKVADLEKEISNLEILIG 60
DB 346 LAKKQTELEKLLDLDPEGKTQDELKAEAEELDKKADELPNKVADLEKEISNLEILIG 405

QY 61 GADSEDDTAALPNKLTAKAELEKTQKELDAALNELGPDGDEE 104
DB 406 GADSEDDTAALQNKLTAKAELEKTQKELDAALNELGPDGDEE 449

RESULT 4
ADM92054
ID ADM92054 standard; protein; 744 AA.
XX
AC ADM92054;
XX
DT 03-JUN-2004 (first entry)
XX
DE S pneumoniae antigenic protein sequence SeqID251.
XX
DE antibacterial; gene therapy; Streptococcus pneumoniae infection;
KW antigenic.
XX
OS Streptococcus pneumoniae.
XX
PN WO2004020609-A2.
XX
PD 11-MAR-2004.
XX
PF 02-SEP-2003; 2003WO-US027401.
XX
PR 30-AUG-2002; 2002US-0407082P.
XX
PA (TUFT ) UNIV TUFTS.
XX
PI Camilli A, Hava DL;
XX
XX WPI; 2004-239189/22.
DR N-PSDB; ADM91817.
XX
XX New Streptococcus pneumoniae nucleic acid molecules, useful for
PT diagnosing, treating and preventing active infections of Streptococcus
PT pneumoniae.
XX
PS Claim 27; SEQ ID NO 251; 123pp; English.
XX
XX This invention relates to novel isolated Streptococcus pneumoniae nucleic
CC acid molecules and the antigenic polypeptides encoded by them. The
CC invention may be useful for the production of compounds with an
CC antibacterial activity or for gene therapy. The nucleic acid molecules,
CC compositions and methods disclosed are useful for treating Streptococcus
CC pneumoniae infection. The present sequence is that of an S pneumoniae
CC protein of the invention.
XX
SQ Sequence 744 AA;
Query Match 96.9%; Score 498; DB 8; Length 744;
Best Local Similarity 98.1%; Pred. No. 3.6e-35;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDLDPEGKTQDELKAEAEELDKKADELPNKVADLEKEISNLEILIG 60
DB 346 LAKKQTELEKLLDLDPEGKTQDELKAEAEELDKKADELPNKVADLEKEISNLEILIG 405

QY 61 GADSEDDTAALPNKLTAKAELEKTQKELDAALNELGPDGDEE 104
DB 406 GADSEDDTAALQNKLTAKAELEKTQKELDAALNELGPDGDEE 449

RESULT 5
AA05590
ID AA05590 standard; protein; 745 AA.
XX
AC AA05590;
XX
DT 24-MAY-2000 (first entry)
XX
DE Streptococcus pneumoniae protein sequence ID301.
XX
XX Streptococcus pneumoniae; vaccine; screening; protein antigen;
KW antibacterial; antiinflammatory; meningitis; infection; diagnosis;
KW pneumococcal disease.
XX
OS Streptococcus pneumoniae.
XX
PN WO200006737-A2.
XX
PD 10-FEB-2000.
XX
PF 27-JUL-1999; 99WO-GB002451.
XX
PR 27-JUL-1998; 98GB-00016337.
PR 19-MAR-1999; 99US-0125164P.
XX
PA (MICR-) MICROBIAL TECHNIQS LTD.
XX
PI Gilbert CFG, Hansbro PM;
XX
XX WPI; 2000-195300/17.
XX
XX New Streptococcal protein, useful as a vaccine, for diagnosis of
PT pneumococcal diseases and for screening agents capable of antagonizing or
PT inhibiting expression of the protein.
XX
PS Claim 2; Page 95; 108pp; English.
XX
XX AA05590 to AA05591 represent specifically claimed protein sequences
CC isolated from Streptococcus pneumoniae. AA05407 to AA05590 represent
CC specifically claimed nucleotide sequences isolated from S. pneumoniae.
CC The sequences have antibacterial and antiinflammatory properties. The
CC protein sequences, and fragments of them, are useful as immunogens and/or
CC antigens. The nucleotide sequences can be used in vaccines and in
CC diagnostic assays. The proteins and nucleotides can be useful for the
CC detection and diagnosis of S. pneumoniae. The protein sequences are also
CC useful for screening an agent capable of antagonising, inhibiting or
CC interfering with the function or expression of the proteins in which the
CC agent is useful for treatment or prophylaxis of S. pneumoniae infection
CC and meningitis. AA05591 to AA05614 represent primers used in the
CC exemplification of the present invention
XX
SQ Sequence 745 AA;
Query Match 96.9%; Score 498; DB 3; Length 745;
Best Local Similarity 98.1%; Pred. No. 3.7e-35;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDLDPEGKTQDELKAEAEELDKKADELPNKVADLEKEISNLEILIG 60
DB 346 LAKKQTELEKLLDLDPEGKTQDELKAEAEELDKKADELPNKVADLEKEISNLEILIG 405

QY 61 GADSEDDTAALPNKLTAKAELEKTQKELDAALNELGPDGDEE 104
DB 406 GADSEDDTAALQNKLTAKAELEKTQKELDAALNELGPDGDEE 449

RESULT 6
AA061217
ID AA061217 standard; protein; 641 AA.
XX
AC AA061217;
XX
```

DT 02-OCT-1998 (first entry)
 XX Streptococcus pneumoniae SP0092 protein.
 DE Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;
 XX detection; pneumonia; otitis media; meningitis.
 KW Streptococcus pneumoniae.
 XX Streptococcus pneumoniae.
 OS
 XX
 XX
 PH Key Location/Qualifiers
 FT Misc-difference 306
 FT /label= unknown
 FT /note= "encoded by NCT"
 XX
 XX WO9818930-A2.
 XX
 XX 07-MAY-1998.
 XX
 XX 30-OCT-1997; 97WO-US019422.
 XX
 XX 31-OCT-1996; 96US-0029960P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Kunsch CA, Choi GH, Johnson LS, Hromockyj A;
 PI
 XX WPI; 1998-272224/24.
 DR N-PSDB; AAV27403.
 DR
 XX Nucleic acid encoding antigenic peptide(s) from Streptococcus pneumoniae
 PT - or their epitope-containing fragments, useful in protective or
 PT therapeutic vaccines, and for diagnosis.
 XX
 XX Claim 11; Page 82; 118pp; English.
 PS
 XX The present sequence represents a protein from Streptococcus pneumoniae.
 CC The nucleic acid sequence encoding the Streptococcus pneumoniae protein
 CC can be useful in vaccines for inducing protective antibodies against
 CC Streptococcus pneumoniae, for treatment or prevention of infection e.g.
 CC pneumonia, otitis media or meningitis. Probes based on the nucleic acid
 CC are used to detect Streptococcus infection (by usual hybridisation or
 CC amplification methods), also for isolating Streptococcus genes or their
 CC allelic variants. The protein can be used similarly to detect specific
 CC antibodies in standard immunoassays, especially for diagnosing or
 CC monitoring infections. Antibodies which bind the protein are used to
 CC detect corresponding antigens, to purify the protein and for passive
 CC immunisation (optionally coupled to a toxin). Vaccines are administered,
 CC e.g. by injection, orally or through the skin, typically at 0.01-1000
 CC (especially 10-300) mu g/ml per dose
 XX
 XX Sequence 641 AA;
 SQ
 Query Match 96.1%; Score 494; DB 2; Length 641;
 Best Local Similarity 97.1%; Pred. No. 6.9e-35;
 Matches 101; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 LAKQTELEKLLDSDPEGKTQDELKAEAEAEADKKADLPNKADELKEISNLEILG 60
 DB 243 LAKQTELEKLLDSDPEGKTQDELKAEAEAEADKKADLPNKADELKEISNLEILG 302
 QY 61 GADSEDDTAALPNKATKKALEKTKQKELDAALNELGPDGDEEE 104
 DB 303 GADXEDDTAALQNKATKKALEKTKQKELDAALNELGPDGDEEE 346
 QY 61 GADSEDDTAALPNKATKKALEKTKQKELDAALNELGPDGDEEE 104
 DB 303 GADXEDDTAALQNKATKKALEKTKQKELDAALNELGPDGDEEE 346
 RESULT 7
 ID ABP54636
 XX ABP54636 standard; protein; 641 AA.
 AC ABP54636;
 XX
 XX 04-SEP-2002 (first entry)
 DT
 XX

DE S. pneumoniae SP092 protein sequence SEQ ID NO:160.
 XX Streptococcus pneumoniae; epitope; vaccine; antigenic protein;
 KW antibacterial; Streptococcal infection; detection.
 XX Streptococcus pneumoniae.
 OS
 XX US2002061545-A1.
 PN
 XX 23-MAY-2002.
 PD
 XX 22-JAN-2001; 2001US-00765272.
 PF
 XX 30-OCT-1997; 97US-00961083.
 PR
 XX (CHOI/) CHOI G H.
 PA (KUNS/) KUNSCH C A.
 PA (BARA/) BARASH S C.
 PA (DILL/) DILLON P J.
 PA (DOUG/) DOUGHERTY B.
 PA (FANN/) FANNON M R.
 PA (ROSE/) ROSEN C A.
 XX
 XX Choi GH, Kunsch CA, Barash SC, Dillon PJ, Dougherty B, Fannon MR;
 PI Rosen CA;
 XX WPI; 2002-479261/51.
 DR N-PSDB; ABQ84871.
 DR
 XX New Streptococcus pneumoniae antigens, useful for detecting Streptococcus
 PT and for preventing or attenuating disease caused by Streptococcus
 PT infection.
 XX
 XX Claim 11; Page 43; 70pp; English.
 PS
 XX ABQ84792 to ABQ84904 represents nucleic acids which encode the
 CC Streptococcus pneumoniae antigens given in ABP54557 to ABP54669. The S.
 CC Streptococcus pneumoniae antigens have antibacterial activity and can be used in
 CC vaccines. The S. pneumoniae antigens can also be used to prevent or
 CC attenuate a Streptococcal infection in an animal. The polynucleotides
 CC encoding the S. pneumoniae antigens can be used to detect Streptococcus
 CC nucleic acids. ABQ84905 to ABQ85130 represent primers used in the cloning
 CC of S. pneumoniae ORFs (open reading frames) which are used in an example
 CC from the present invention
 XX
 XX Sequence 641 AA;
 SQ
 Query Match 96.1%; Score 494; DB 5; Length 641;
 Best Local Similarity 97.1%; Pred. No. 6.9e-35;
 Matches 101; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 LAKQTELEKLLDSDPEGKTQDELKAEAEAEADKKADLPNKADELKEISNLEILG 60
 DB 243 LAKQTELEKLLDSDPEGKTQDELKAEAEAEADKKADLPNKADELKEISNLEILG 302
 QY 61 GADSEDDTAALPNKATKKALEKTKQKELDAALNELGPDGDEEE 104
 DB 303 GADXEDDTAALQNKATKKALEKTKQKELDAALNELGPDGDEEE 346
 RESULT 8
 ID ADC45241
 XX ADC45241 standard; protein; 641 AA.
 AC ADC45241;
 XX
 XX 18-DEC-2003 (first entry)
 DT
 XX S. pneumoniae antigenic protein SP092.
 DE
 XX Antigen; bacterial infection; vaccine; pneumonia; antibacterial.
 KW
 XX Streptococcus pneumoniae.
 XX


```

XX US6573082-B1.
PN
XX
XX
XX 03-JUN-2003.
XX
XX 28-MAR-2000; 2000US-00536784.
XX
XX 31-OCT-1996; 96US-0029960P.
PR 30-OCT-1997; 97US-00961083.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Choi GH, Kunsch CA, Barash SC, Dillon PJ, Dougherty B, Fannon MR;
PI Rosen CA;
PI
XX WPI; 2003-764574/72.
DR N-PSDB; ADC45240.
DR
XX Novel polynucleotide encoding Streptococcus pneumoniae polypeptides
PT useful for producing vaccines for prevention or attenuation of infection
PT by Streptococcus pneumoniae.
XX
PS Example 1; SEQ ID NO 160; 58pp; English.
XX
CC The invention relates to an isolated polynucleotide consisting of a
CC Streptococcus pneumoniae nucleic acid (appearing as ADC45122 and encoding
CC SP028) one of 113 disclosed nucleic acids encoding 113 S. pneumoniae
CC antigens. Also included are making a recombinant vector by inserting the
CC nucleic acid into a vector, an isolated polynucleotide consisting of at
CC least 50 or 100 contiguous nucleotides of the SP028 nucleic acid, and a
CC recombinant host cell comprising the SP028 polynucleotide. The nucleic
CC acids are useful as DNA vaccine against Streptococcus pneumoniae
CC infection (e.g. pneumonia). Nucleic acids derived from the S. pneumoniae
CC antigen nucleic acids are useful as probes for use in diagnostic methods
CC for detecting S. pneumoniae gene expression. The present sequence
CC represents an S. pneumoniae antigenic protein.
XX
SQ Sequence 641 AA;
Query Match 96.1%; Score 494; DB 7; Length 641;
Best Local Similarity 97.1%; Pred. No. 6.9e-35;
Matches 101; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 LAKKQTELEKLLDLPDGKQTQDELDKAEAEALDKKADLPKNKVLADLEKEISNLSILLG 60
DB 243 LAKKQTELEKLLDLPDGKQTQDELDKAEAEALDKKADLPKNKVLADLEKEISNLSILLG 302
QY 61 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 104
DB 303 GADXEDDTAALQNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 346
RESULT 9
ABW02601
ID ABW02601 standard; protein; 213 AA.
XX
XX ABW02601;
AC
XX
XX 12-FEB-2004 (first entry)
DT
XX
DE Bg8090c pneumococcal surface protein A (PspA) central region.
DE
XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
XX immunological; gene therapy; immunostimulant.
XX
XX Unidentified.
OS
XX
XX Key Location/Qualifiers
FH Key Location/Qualifiers
FT Misc-difference 2
FT /label= Unknown
XX
XX US6592876-B1.
PN
XX

```

```

PD 15-JUL-2003.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX 20-APR-1993; 93US-00048896.
PR 06-JUN-1995; 95US-00465746.
XX
XX (UABR-) UAB RES FOUND.
PA
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
PI WPI; 2003-862841/80.
DR
XX Immunological composition for obtaining expression products used for
PT detecting the presence of Streptococcus pneumoniae or its strain,
PT comprises at least two different full length isolated gene encoding
PT pneumococcal surface protein A.
XX
PS Example 6; SEQ ID NO 47; 121pp; English.
XX
CC The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspAs) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antigenic,
CC immunological or vaccine compositions, for eliciting antibodies, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Bg8090c
CC pneumococcal surface protein A (PspA) central region. This sequence is
CC used in the exemplification of the invention
XX
SQ Sequence 213 AA;
Query Match 95.3%; Score 490; DB 7; Length 213;
Best Local Similarity 95.2%; Pred. No. 4.3e-35;
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 LAKKQTELEKLLDLPDGKQTQDELDKAEAEALDKKADLPKNKVLADLEKEISNLSILLG 60
DB 59 LAKKQTELEKLLDLPDGKQTQDELDKAEAEALDKKADLPKNKVLADLEKEISNLSILLG 118
QY 61 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 104
DB 119 GADPDDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 162
RESULT 10
ABU08487
ID ABU08487 standard; protein; 8991 AA.
XX
XX ABU08487;
AC
XX
XX 24-JUN-2003 (first entry)
DT
XX
DE S. pneumoniae pneumococcal surface protein A (PspA) protein.
DE
XX Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
XX alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
XX antibacterial.
XX
XX Streptococcus pneumoniae.
OS
XX
XX Key Location/Qualifiers
FH Key Location/Qualifiers
FT Misc-difference 1.8991
FT /note= "All Xaa residues within this sequence are
FT unknown"
XX
XX US6500613-B1.
PN
XX

```

```

PD 31-DEC-2002.
XX
XX PF 16-SEP-1996; 96US-00714741.
XX
XX PR 15-SEP-1995; 95US-00529055.
XX
XX PA (UYAL-) UNIV ALABAMA.
XX
XX PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
XX PI Hollingshead S, Tart R, Brooks-Walter A;
XX
XX DR WPI; 2003-361534/34.
XX
XX PT Isolated PspC amino acid sequence used as polymerase chain reaction or
XX PT hybridization probe, comprises pneumococcal surface protein having alpha-
XX PT helical, proline rich and repeat regions.
XX
XX PS Disclosure; Col 145-188; 186pp; English.
XX
XX CC The present invention relates to the isolation of Streptococcus
XX CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
XX CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-
XX CC like protein having alpha-helical, proline rich and repeat regions. The
XX CC PspC and PspA proteins may be used in a vaccine to protect against
XX CC pneumococcal infections. The polynucleotide sequences encoding PspC and
XX CC PspA may be used for the expression of the proteins, and as PCR primers
XX CC or hybridisation probes. The present sequence represents S. pneumoniae
XX CC PspA protein
XX
XX SQ Sequence 8991 AA;

Query Match 95.3%; Score 490; DB 6; Length 8991;
Best Local Similarity 95.2%; Pred. No. 3.2e-33;
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKQTELEKLLDLSLDPGKTKQDELKAEAEALDKKADLPNPKVADLEKEISNLEILLG 60
DB 4625 LAKQTELEKLLDNLDPGKTKQDELKAEAEALDKKADLPNPKVADLEKEISNLEILLG 4684
QY 61 GADSEDDTAALPNKLTAKAELEKTKOKELDAALNELGPDGDEEE 104
DB 4685 GADPEDDTAALPNKLTAKAEFEKTFKELDAALNELGPDGDEEE 4728

RESULT 11
AAW14567
ID AAW14567 standard; protein; 213 AA.
XX
XX AC AAW14567;
XX
XX DT 17-OCT-2003 (revised)
XX DT 28-OCT-1997 (first entry)
XX
XX DE Streptococcus pneumoniae PspA central region.
XX
XX KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
XX KW bacteraemia; pneumonia.
XX
XX OS Streptococcus pneumoniae; strain Bg8090.
XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 2
XX FT /note= "unidentified amino acid"
XX
XX PN W09709994-A1.
XX
XX PD 20-MAR-1997.
XX
XX PF 16-SEP-1996; 96WO-US014819.
XX
XX PR 15-SEP-1995; 95US-00529055.
XX
XX PA (UABR-) UAB RES FOUND.

XX
XX PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
XX PI Hollingshead S, Tart R, Brooks-Walter A;
XX
XX DR WPI; 2003-862841/80.
XX
XX PT Immunological composition for obtaining expression products used for
XX PT detecting the presence of Streptococcus pneumoniae or its strain,
XX PT comprises at least two different full length isolated gene encoding

```

```

XX
XX PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
XX PI Hollingshead S, Tart R, Brooks-Walter A;
XX
XX DR WPI; 1997-202002/18.
XX
XX PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used
XX PT in vaccines for protecting animals against S.pneumoniae infection.
XX
XX PS Example 6; Fig 13; 296pp; English.
XX
XX CC This sequence shows the central portion, including the C-terminus of the
XX CC alpha-helix region and some of the proline-rich region, of pneumococcal
XX CC surface protein A (PspA) of Streptococcus pneumoniae strain Bg8090.
XX CC Comparison of the N-terminal and central regions (AAW14533-57 and
XX CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
XX CC be used to divide the strains into several families based on sequence
XX CC homologies. PspA polypeptides, or fragments of them, can be used in
XX CC vaccines to protect animals against S. pneumoniae infection and hence for
XX CC the prevention of diseases such as otitis media, meningitis, bacteraemia
XX CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
XX CC region and the immediate 5' tip of the coding sequence are likely to be
XX CC the critical sequences for predicting PspA cross-reactions and vaccine
XX CC composition. (Updated on 17-Oct-2003 to standardise OS field)
XX
XX SQ Sequence 213 AA;

Query Match 94.9%; Score 488; DB 2; Length 213;
Best Local Similarity 95.2%; Pred. No. 6.5e-35;
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKQTELEKLLDLSLDPGKTKQDELKAEAEALDKKADLPNPKVADLEKEISNLEILLG 60
DB 59 LAKQTELEKLLDNLDPGKTKQDELKAEAEALDKKADLPNPKVADLEKEISNLEILLG 118
QY 61 GADSEDDTAALPNKLTAKAELEKTKOKELDAALNELGPDGDEEE 104
DB 119 GADPEDDTAALPNKLTAKAEFEKTFKELDAALNELGPDGDEEE 162

RESULT 12
ABW02598
ID ABW02598 standard; protein; 197 AA.
XX
XX AC ABW02598;
XX
XX DT 12-FEB-2004 (first entry)
XX
XX DE Ac122c pneumococcal surface protein A (PspA) central region.
XX
XX KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
XX KW immunological; gene therapy; immunostimulant.
XX
XX OS Unidentified.
XX
XX PN US6592876-B1.
XX
XX PD 15-JUL-2003.
XX
XX PF 15-SEP-1995; 95US-00529055.
XX
XX PR 20-APR-1993; 93US-00048896.
XX PR 06-JUN-1995; 95US-00465746.
XX
XX PA (UABR-) UAB RES FOUND.
XX
XX PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX PI WPI; 2003-862841/80.
XX
XX PT Immunological composition for obtaining expression products used for
XX PT detecting the presence of Streptococcus pneumoniae or its strain,
XX PT comprises at least two different full length isolated gene encoding

```

```

PT pneumococcal surface protein A.
PS Example 6; SEQ ID NO 44; 121pp; English.
XX
XX The present invention relates to an immunological composition comprising
XX at least 2 different full length isolated genes encoding pneumococcal
XX surface protein A (PspA) from different groups based on restriction
XX fragment polymorphism analysis. The invention is useful for obtaining
XX expression products by recombinant techniques to detect, determine,
XX isolate or diagnose the presence of Streptococcus pneumoniae or its
XX strain. The expression product is useful for preparing antibodies, an
XX immunological or vaccine compositions, for eliciting antigenic, an
XX immunological response (other than or additional to antibodies) or a
XX protective response (including antibody or other immunological response
XX by administering compositions to a host). The invention is also useful as
XX vaccines and in gene therapy. The present sequence is Acl22c pneumococcal
XX surface protein A (PspA) central region. This sequence is used in the
XX exemplification of the invention
XX
SQ Sequence 197 AA;
Query Match 94.6%; Score 486; DB 7; Length 197;
Best Local Similarity 95.2%; Pred. No. 8.9e-35;
Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 LAKKQTELEKLLDSLDPGKTDQELDKAEAEAEELDKKADLPNKVADLEKEISNLEILG 60
DB 22 LAQKQTELEKLLDSLDPGKTDQELDKAEAEAEELDKKADLPNKVADLEKEISNLEILG 81
QY 61 GADSEDDTAALPNKLTAKAELEKTKQELDAALNELGPDGDEE 104
DB 82 GADSEDDTAALPNKLTAKAELEKTKQELDAALNELGPDGDEE 125

RESULT 13
ID ABW02606 standard; protein; 233 AA.
XX
XX ABW02606;
XX
XX 12-FEB-2004 (first entry)
XX
XX Bfl019c pneumococcal surface protein A (PspA) central region.
XX
XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
XX immunological; gene therapy; immunostimulant.
XX
XX Unidentified.
XX
XX Key Location/Qualifiers
FH Misc-difference 1..233
FT /note= "Xaa = Unknown amino acid"
XX
XX US6592876-B1.
XX
XX 15-JUL-2003.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX 20-APR-1993; 93US-00048896.
XX
XX 06-JUN-1995; 95US-00465746.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX WPI; 2003-862841/80.
XX
XX Immunological composition for obtaining expression products used for
XX detecting the presence of Streptococcus pneumoniae or its strain,
XX comprises at least two different full length isolated gene encoding
XX pneumococcal surface protein A.

```

```

PS Example 6; SEQ ID NO 52; 121pp; English.
XX
XX The present invention relates to an immunological composition comprising
XX at least 2 different full length isolated genes encoding pneumococcal
XX surface protein A (PspA) from different groups based on restriction
XX fragment polymorphism analysis. The invention is useful for obtaining
XX expression products by recombinant techniques to detect, determine,
XX isolate or diagnose the presence of Streptococcus pneumoniae or its
XX strain. The expression product is useful for preparing antibodies, an
XX immunological or vaccine compositions, for eliciting antigenic, an
XX immunological response (other than or additional to antibodies) or a
XX protective response (including antibody or other immunological response
XX by administering compositions to a host). The invention is also useful as
XX vaccines and in gene therapy. The present sequence is Efl019c
XX pneumococcal surface protein A (PspA) central region. This sequence is
XX used in the exemplification of the invention
XX
SQ Sequence 233 AA;
Query Match 90.5%; Score 465; DB 7; Length 233;
Best Local Similarity 92.3%; Pred. No. 7.5e-33;
Matches 96; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 1 LAKKQTELEKLLDSLDPGKTDQELDKAEAEAEELDKKADLPNKVADLEKEISNLEILG 60
DB 53 LAQKQTELEKLLDSLDPGKTDQELDKAEAEAEELDKKADLPNKVADLEKEISNLEILG 112
QY 61 GADSEDDTAALPNKLTAKAELEKTKQELDAALNELGPDGDEE 104
DB 113 GADSEDDTAALPNKLTAKAELEKTKQELDAALNELGPDGDEE 156

RESULT 14
AAW14564
ID AAW14564 standard; protein; 196 AA.
XX
XX AAW14564;
XX
XX 17-OCT-2003 (revised)
XX 28-OCT-1997 (first entry)
XX
XX Streptococcus pneumoniae PspA central region.
XX
XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
XX bacteraemia; pneumonia.
XX
XX Streptococcus pneumoniae; strain Acl22.
XX
XX WO9709994-A1.
XX
XX 20-MAR-1997.
XX
XX 16-SEP-1996; 96WO-US014819.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
XX Hollingshead S, Tart R, Brooks-Walter A;
XX WPI; 1997-202002/18.
XX
XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
XX in vaccines for protecting animals against S.pneumoniae infection.
XX
XX Example 6; Fig 13; 296pp; English.
XX
XX This sequence shows the central portion, including the C-terminus of the
XX alpha-helix region and some of the proline-rich region, of pneumococcal
XX surface protein A (PspA) of Streptococcus pneumoniae strain Acl22.
XX Comparison of the N-terminal and central regions (AAW14533-57 and
XX AAW14562-91) of PspA polypeptides from different pneumococcal strains can

```


Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	514	100.0	104	2	US-08-710-749-19		Sequence 19, Appl
2	514	100.0	104	4	US-09-147-875A-20		Sequence 20, Appl
3	500	97.3	104	4	US-09-147-875A-21		Sequence 21, Appl
4	494	96.1	104	2	US-08-710-749-20		Sequence 20, Appl
5	494	96.1	641	3	US-08-961-083-160		Sequence 160, Appl
6	494	96.1	641	4	US-09-536-784-160		Sequence 160, Appl
7	490	95.3	213	4	US-08-529-055-47		Sequence 47, Appl
8	490	95.3	8991	4	US-08-714-741-32		Sequence 32, Appl
9	486	94.6	197	4	US-08-529-055-44		Sequence 44, Appl
10	484	94.2	102	2	US-08-710-749-21		Sequence 21, Appl
11	484	94.2	102	4	US-09-147-875A-18		Sequence 18, Appl
12	465	90.5	233	4	US-08-529-055-52		Sequence 52, Appl
13	378	73.5	80	2	US-08-710-749-18		Sequence 18, Appl
14	378	73.5	80	4	US-09-147-875A-19		Sequence 19, Appl
15	319	62.1	108	2	US-08-710-749-26		Sequence 26, Appl
16	319	62.1	108	4	US-09-147-875A-23		Sequence 23, Appl
17	319	62.1	211	4	US-08-529-055-67		Sequence 67, Appl
18	313	60.9	108	2	US-08-710-749-24		Sequence 24, Appl
19	313	60.9	108	4	US-09-147-875A-25		Sequence 25, Appl
20	313	60.9	232	4	US-08-529-055-70		Sequence 70, Appl
21	313	60.9	458	4	US-08-529-055-73		Sequence 73, Appl
22	311	60.5	108	4	US-09-147-875A-24		Sequence 24, Appl
23	309	60.1	106	4	US-09-147-875A-22		Sequence 22, Appl
24	304	59.1	212	4	US-08-529-055-68		Sequence 68, Appl
25	301	58.6	108	2	US-08-710-749-22		Sequence 22, Appl
26	301	58.6	108	2	US-08-710-749-23		Sequence 23, Appl
27	263	51.2	108	2	US-08-710-749-25		Sequence 25, Appl

```
Qy 61 GADSEDDTAALPNKLTAKAELEKTKQELDAALNELGPDGDEE 104
|||||
Db 61 GADSEDDTAALPNKLTAKAELEKTKQELDAALNELGPDGDEE 104

RESULT 2
US-09-147-875A-20
; Sequence 20, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 20
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-20

Query Match 100.0%; Score 514; DB 4; Length 104;
Best Local Similarity 100.0%; Pred. No. 2.2e-42;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LAKQTELEKLLDSLDPEGKTQDELKAEAEALDKKADELPNKVADLEKEISNLEILLG 60
|||||
Db 1 LAKQTELEKLLDSLDPEGKTQDELKAEAEALDKKADELPNKVADLEKEISNLEILLG 60

Qy 61 GADSEDDTAALPNKLTAKAELEKTKQELDAALNELGPDGDEE 104
|||||
Db 61 GADSEDDTAALPNKLTAKAELEKTKQELDAALNELGPDGDEE 104

RESULT 3
US-09-147-875A-21
; Sequence 21, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 21
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-21

Query Match 97.3%; Score 500; DB 4; Length 104;
Best Local Similarity 97.1%; Pred. No. 4.8e-41;
Matches 101; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LAKQTELEKLLDSLDPEGKTQDELKAEAEALDKKADELPNKVADLEKEISNLEILLG 60
|||||
Db 1 LAKQTELEKLLDNLDPGKTQDELKAEAEALDKKADELPNKVADLEKEISNLEILLG 60

Qy 61 GADSEDDTAALPNKLTAKAELEKTKQELDAALNELGPDGDEE 104
|||||
Db 61 GADPEDTAALPNKLTAKAELEKTKQELDAALNELGPDGDEE 104

RESULT 4
US-08-710-749-20
; Sequence 20, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
```

```
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-20

Query Match 96.1%; Score 494; DB 2; Length 104;
Best Local Similarity 96.2%; Pred. No. 1.8e-40;
Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LAKQTELEKLLDSLDPEGKTQDELKAEAEALDKKADELPNKVADLEKEISNLEILLG 60
|||||
Db 1 LAKQTELEKLLDNLDPGKTQDELKAEAEALDKKADELPNKVADLEKEISNLEILLG 60

Qy 61 GADSEDDTAALPNKLTAKAELEKTKQELDAALNELGPDGDEE 104
|||||
Db 61 GADPEDDTAALPNKLTAKAELEKTKQELDAALNELGPDGDEE 104

RESULT 5
US-08-961-083-160
; Sequence 160, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
```

```
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Brookes, A. Anders
;; REGISTRATION NUMBER: 36,373
;; REFERENCE/DOCKET NUMBER: PB340P2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (301) 309-8504
;; TELEFAX: (301) 309-8512
;; INFORMATION FOR SEQ ID NO: 160:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 641 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-961-083-160

Query Match 96.1%; Score 494; DB 3; Length 641;
Best Local Similarity 97.1%; Pred. No. 1.7e-39;
Matches 101; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDLSDPEGKTQDELDKAEAEELDKKADELPKNKVDLEKEISNLEILG 60
Db 243 LAKKQTELEKLDLSDPEGKTQDELDKAEAEELDKKADELPKNKVDLEKEISNLEILG 302
QY 61 GADSEDDTAALPNKLTAKAELEKTOKELDAALNELGPDGDEE 104
Db 303 GADXEDDTAALQNKLTAKAELEKTOKELDAALNELGPDGDEE 346

RESULT 6
US-09-536-784-160
; Sequence 160, Application US/09536784
; Patent No. 6573082
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/536,784
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB340P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 160:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; SEQUENCE DESCRIPTION: SEQ ID NO: 160:
US-09-536-784-160

Query Match 96.1%; Score 494; DB 4; Length 641;
Best Local Similarity 97.1%; Pred. No. 1.7e-39;
Matches 101; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDLSDPEGKTQDELDKAEAEELDKKADELPKNKVDLEKEISNLEILG 60
Db 243 LAKKQTELEKLDLSDPEGKTQDELDKAEAEELDKKADELPKNKVDLEKEISNLEILG 302
QY 61 GADSEDDTAALPNKLTAKAELEKTOKELDAALNELGPDGDEE 104
Db 303 GADXEDDTAALQNKLTAKAELEKTOKELDAALNELGPDGDEE 346

RESULT 7
US-08-529-055-47
; Sequence 47, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 213 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-47

Query Match 95.3%; Score 490; DB 4; Length 213;
Best Local Similarity 95.2%; Pred. No. 1.1e-39;
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDLSDPEGKTQDELDKAEAEELDKKADELPKNKVDLEKEISNLEILG 60
Db 59 LAKKQTELEKLDLSDPEGKTQDELDKAEAEELDKKADELPKNKVDLEKEISNLEILG 118
```


;/ MEDIUM TYPE: Floppy disk
;/ COMPUTER: IBM PC compatible
;/ OPERATING SYSTEM: PC-DOS/MS-DOS
;/ SOFTWARE: Patent In Release #1.0, Version #1.30
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/09/710,749
;/ FILING DATE: 20-SEP-1996
;/ CLASSIFICATION: 435
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Frommer, William S.
;/ REGISTRATION NUMBER: 25,506
;/ REFERENCE/DOCKET NUMBER: 454312-2074
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (212) 840-3333
;/ TELEFAX: (212) 840-0712
;/ INFORMATION FOR SEQ ID NO: 21:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 102 amino acids
;/ TYPE: amino acid
;/ STRANDEDNESS: n/a
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: amino acid
;/ US-08-710-749-21

Query Match 94.2%; Score 484; DB 2; Length 102;
Best Local Similarity 98.1%; Pred. No. 1.6e-39;
Matches 102; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLDSLDPEGKTQDELKAEAEELDKKADLPKVKADLKEISNLEILLG 60
Db 1 LAKKQTELEKLLD-LDPEGKTQDELKAEAEELDKKADLPKVKADLKEISNLEILLG 58
QY 61 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 104
Db 59 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 102

RESULT 11
US-09-147-875A-18
; Sequence 18, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 18
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-147-875A-18

Query Match 94.2%; Score 484; DB 4; Length 102;
Best Local Similarity 98.1%; Pred. No. 1.6e-39;
Matches 102; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLDSLDPEGKTQDELKAEAEELDKKADLPKVKADLKEISNLEILLG 60
Db 1 LAKKQTELEKLLD-LDPEGKTQDELKAEAEELDKKADLPKVKADLKEISNLEILLG 58
QY 61 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 104
Db 59 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 102

RESULT 12
US-08-529-055-52
; Sequence 52, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:

;/ APPLICANT: Briles, David E.
;/ APPLICANT: McDaniel, Larry S.
;/ APPLICANT: Swiatlo, Edwin
;/ APPLICANT: Yother, Janet
;/ APPLICANT: Brooks-Walter, Alexis
;/ TITLE OF INVENTION: Pneumococcal Genes, Portions
;/ TITLE OF INVENTION: Thereof, Expression Products
;/ TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
;/ NUMBER OF SEQUENCES: 73
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: Curtis, Morris & Safford, P.C.
;/ STREET: 530 Fifth Avenue
;/ CITY: New York
;/ STATE: NY
;/ COUNTRY: USA
;/ ZIP: 10036
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Floppy disk
;/ COMPUTER: IBM PC compatible
;/ OPERATING SYSTEM: PC-DOS/MS-DOS
;/ SOFTWARE: Patent In Release #1.0, Version #1.30
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/529,055
;/ FILING DATE: 15-SEP-1995
;/ CLASSIFICATION: 435
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Frommer, William S.
;/ REGISTRATION NUMBER: 25,506
;/ REFERENCE/DOCKET NUMBER: 454312-2400
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (212) 840-3333
;/ TELEFAX: (212) 840-0712
;/ INFORMATION FOR SEQ ID NO: 52:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 233 amino acids
;/ TYPE: amino acid
;/ STRANDEDNESS: single
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: peptide
;/ US-08-529-055-52

Query Match 90.5%; Score 465; DB 4; Length 233;
Best Local Similarity 92.3%; Pred. No. 3e-37;
Matches 96; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDSLDPEGKTQDELKAEAEELDKKADLPKVKADLKEISNLEILLG 60
Db 53 LAKKQTELEKLLDSLDPEGKTQDELKAEAEELDKKADLPKVKADLKEISNLEILLG 112
QY 61 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 104
Db 113 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 156

RESULT 13
US-08-710-749-18
; Sequence 18, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/710,749
; APPLICATION NUMBER: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-710-749-18

Query Match 73.5%; Score 378; DB 2; Length 80;
Best Local Similarity 96.2%; Pred. No. 1.9e-29;
Matches 77; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 25 LDKAEAEALDKKADLPNKVADLEKEISNLEILLGADSEDDTAALPNKLTATKKALEK 84
Db 1 LDKAEAEALDKKADLPNKVADLEKEISNLEILLGADSEDDTAALPNKLTATKKALEK 60
QY 85 TOKELDAALNELGPDGDEE 104
Db 61 TOKELDAALNELGPDGDEE 80

RESULT 14
US-09-147-875A-19
; Sequence 19, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-147-875A-19

Query Match 73.5%; Score 378; DB 4; Length 80;
Best Local Similarity 96.2%; Pred. No. 1.9e-29;
Matches 77; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 25 LDKAEAEALDKKADLPNKVADLEKEISNLEILLGADSEDDTAALPNKLTATKKALEK 84
Db 1 LDKAEAEALDKKADLPNKVADLEKEISNLEILLGADSEDDTAALPNKLTATKKALEK 60
QY 85 TOKELDAALNELGPDGDEE 104
Db 61 TOKELDAALNELGPDGDEE 80

RESULT 15
US-08-710-749-26
; Sequence 26, Application US/08710749
; Patent No. 5955089
```

```
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 108 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-710-749-26
```

```
Query Match 62.1%; Score 319; DB 2; Length 108;
Best Local Similarity 64.8%; Pred. No. 1.3e-23;
Matches 70; Conservative 13; Mismatches 21; Indels 4; Gaps 2;

QY 1 LAKKQTELEKLLDLPDGKTDQELDKAEAEALDKKADLPNKVADLEKEISNLEILLG 60
Db 1 LKAEAELENLLSTLDPEGKTQDELDKAEAEALDKKADLPNKVADLEKEISNLEILLG 60
QY 61 GADS---ED-DTAALPNKLTATKKALEKTQKELDAALNELGPDGDEE 104
Db 61 DAETNNVEDYIKEGLEAEATKQAELEKTPKELDAALNELGPDGDEE 108
```

Search completed: June 21, 2005, 10:25:23
Job time : 19.4204 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 10:12:15 ; Search time 66.4327 Seconds
(without alignments)
601.118 Million cell updates/sec

Title: US-10-674-755-20

Perfect score: 514

Sequence: 1 LAKKQTELEKLLDLDPEK.....TKELDAALNELGPDGDEE 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	514	100.0	104	15	US-10-674-755-20
2	500	97.3	104	15	US-10-674-755-21
3	498	96.9	744	10	US-09-769-787-184
4	498	96.9	744	17	US-10-472-928-32
5	494	96.1	641	9	US-09-765-272-160
6	490	95.3	213	15	US-10-299-636-62
7	486	94.6	197	15	US-10-299-636-59
8	484	94.2	102	15	US-10-674-755-18
9	465	90.5	233	15	US-10-299-638-67
10	378	73.5	80	15	US-10-674-755-19
11	319	62.1	108	15	US-10-674-755-23

12	319	62.1	211	15	US-10-299-636-82	Sequence 82, Appl
13	313	60.9	108	15	US-10-674-755-25	Sequence 25, Appl
14	313	60.9	232	15	US-10-299-636-85	Sequence 85, Appl
15	313	60.9	275	16	US-10-414-532-1	Sequence 1, Appl
16	313	60.9	458	15	US-10-299-636-88	Sequence 88, Appl
17	313	60.9	653	16	US-10-414-532-26	Sequence 26, Appl
18	311	60.5	108	15	US-10-674-755-24	Sequence 24, Appl
19	309	60.1	106	15	US-10-674-755-22	Sequence 22, Appl
20	304	59.1	212	15	US-10-299-636-83	Sequence 83, Appl
21	277	53.9	459	16	US-10-702-305A-18	Sequence 18, Appl
22	263	51.2	108	15	US-10-674-755-26	Sequence 26, Appl
23	256	49.8	185	15	US-10-299-636-84	Sequence 84, Appl
24	198	38.5	487	16	US-10-414-532-34	Sequence 34, Appl
25	198	38.5	487	16	US-10-414-532-21	Sequence 21, Appl
26	198	38.5	524	16	US-10-414-532-28	Sequence 28, Appl
27	193.5	37.6	119	15	US-10-674-755-27	Sequence 27, Appl
28	193.5	37.6	215	15	US-10-299-636-58	Sequence 58, Appl
29	185.5	36.1	290	16	US-10-414-532-65	Sequence 65, Appl
30	180	35.0	230	16	US-10-414-532-32	Sequence 32, Appl
31	180	35.0	230	16	US-10-414-532-19	Sequence 19, Appl
32	128	24.9	354	15	US-10-299-636-105	Sequence 105, Appl
33	128	24.9	588	15	US-10-299-636-96	Sequence 96, Appl
34	128	24.9	619	10	US-09-882-774-1	Sequence 1, Appl
35	128	24.9	619	15	US-10-282-122A-73702	Sequence 73702, A
36	128	24.9	619	16	US-10-414-532-72	Sequence 72, Appl
37	126	24.5	141	14	US-10-254-995-2	Sequence 2, Appl
38	126	24.5	204	15	US-10-299-636-66	Sequence 66, Appl
39	126	24.5	589	9	US-09-748-875-14	Sequence 14, Appl
40	126	24.5	589	10	US-09-298-523B-14	Sequence 14, Appl
41	126	24.5	589	15	US-10-299-636-97	Sequence 97, Appl
42	126	24.5	643	15	US-10-299-636-95	Sequence 95, Appl
43	126	24.5	670	9	US-09-748-875-63	Sequence 63, Appl
44	126	24.5	670	10	US-09-298-523B-63	Sequence 63, Appl
45	126	24.5	690	9	US-09-748-875-61	Sequence 61, Appl

ALIGNMENTS

RESULT 1

US-10-674-755-20

; Sequence 20, Application US/10674755

; Publication No. US20040067237A1

; GENERAL INFORMATION:

; APPLICANT: BECKER et al.

; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS

; FILE REFERENCE: 454312-2471

; CURRENT APPLICATION NUMBER: US/10/674,755

; PRIOR FILING DATE: 2003-09-30

; PRIOR APPLICATION NUMBER: US/09/147,875A

; PRIOR FILING DATE: 1999-05-24

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 20

; LENGTH: 104

; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae

; US-10-674-755-20

Query Match 100.0%; Score 514; DB 15; Length 104;
Best Local Similarity 100.0%; Pred. No. 8.8e-36;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDLDPEKQTQDELKKEAEELDKKADLPNKVADLEKEISNLEILLG 60

pb 1 LAKKQTELEKLLDLDPEKQTQDELKKEAEELDKKADLPNKVADLEKEISNLEILLG 60

QY 61 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 104

Db 61 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 104

RESULT 2

```
US-10-674-755-21
; Sequence 21, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 104
; TYPE: PR1
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-21

Query Match          97.3%; Score 500; DB 15; Length 104;
Best Local Similarity 97.1%; Pred. No. 1.3e-34;
Matches 101; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LAKQTELEKLLSLDPEGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 60
Db 1 LAKQTELEKLLDNLDPGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 60

Qy 61 GADSEDDTAALPNKLTAKAELEKTQKELDAALNELGPDGDEEE 104
Db 61 GADPEDTAALPNKLTAKAELEKTQKELDAALNELGPDGDEEE 104

RESULT 3
US-09-769-787-184
; Sequence 184, Application US/09769787
; Publication No. US20030091577A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 184
; LENGTH: 744
; TYPE: PR1
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-184

Query Match          96.9%; Score 498; DB 10; Length 744;
Best Local Similarity 98.1%; Pred. No. 2e-33;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LAKQTELEKLLSLDPEGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 60
Db 346 LAKQTELEKLLSLDPEGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 405

Qy 61 GADSEDDTAALPNKLTAKAELEKTQKELDAALNELGPDGDEEE 104
Db 406 GADSEDDTAALPNKLTAKAELEKTQKELDAALNELGPDGDEEE 449

RESULT 4
US-10-472-928-32
; Sequence 32, Application US/10472928
; Publication No. US20050020813A1
```

```
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 32
; LENGTH: 744
; TYPE: PR1
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: pneumococcal surface protein A (pspA)
; OTHER INFORMATION: Cellular location: outside
; OTHER INFORMATION: Feature of note: WY motif
; OTHER INFORMATION: Similar to strain R6 sequence 15902165 (e-179)
US-10-472-928-32

Query Match          96.9%; Score 498; DB 17; Length 744;
Best Local Similarity 98.1%; Pred. No. 2e-33;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LAKQTELEKLLSLDPEGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 60
Db 346 LAKQTELEKLLSLDPEGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 405

Qy 61 GADSEDDTAALPNKLTAKAELEKTQKELDAALNELGPDGDEEE 104
Db 406 GADSEDDTAALPNKLTAKAELEKTQKELDAALNELGPDGDEEE 449

RESULT 5
US-09-765-272-160
; Sequence 160, Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 160:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids
; TYPE: amino acid
```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 160:
US-09-765-272-160

Query Match          96.1%; Score 494; DB 9; Length 641;
Best Local Similarity 97.1%; Pred. No. 3.6e-33;
Matches 101; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDLPDGKTDQELDKAEAEAEADKKADELPNKVADLEKEISNLEILLG 60
   |||||
Db 243 LAKKQTELEKLLDLPDGKTDQELDKAEAEAEADKKADELPNKVADLEKEISNLEILLG 302
   |||||

QY 61 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 104
   |||||
Db 303 GADXEDDTAALQNLATKKAELKTKQKELDAALNELGPDGDEEE 346
   |||||

RESULT 6
US-10-299-636-62
; Sequence 62, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; PRIOR FILING DATE: 2002-11-19
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (2)
; OTHER INFORMATION: Xaa at position 2 is unknown
US-10-299-636-62

Query Match          95.3%; Score 490; DB 15; Length 213;
Best Local Similarity 95.2%; Pred. No. 2.1e-33;
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDLPDGKTDQELDKAEAEAEADKKADELPNKVADLEKEISNLEILLG 60
   |||||
Db 59 LAKKQTELEKLLDLPDGKTDQELDKAEAEAEADKKADELPNKVADLEKEISNLEILLG 118
   |||||

QY 61 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 104
   |||||
Db 119 GADPEDDTAALPNKLTAKKAEFEKTKPELDAALNELGPDGDEEE 162
   |||||

RESULT 7
US-10-299-636-59
; Sequence 59, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
```

```
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-59

Query Match          94.6%; Score 486; DB 15; Length 197;
Best Local Similarity 95.2%; Pred. No. 4.2e-33;
Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDLPDGKTDQELDKAEAEAEADKKADELPNKVADLEKEISNLEILLG 60
   |||||
Db 22 LAQKQTELEKLLDLPDGKTDQELDKAEAEAEADKKADELPNKVADLEKEISNLEILLG 81
   |||||

QY 61 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 104
   |||||
Db 82 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 125
   |||||

RESULT 8
US-10-674-755-18
; Sequence 18, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-18

Query Match          94.2%; Score 484; DB 15; Length 102;
Best Local Similarity 98.1%; Pred. No. 2.8e-33;
Matches 102; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLDLPDGKTDQELDKAEAEAEADKKADELPNKVADLEKEISNLEILLG 60
   |||||
Db 1 LAKKQTELEKLLDLPDGKTDQELDKAEAEAEADKKADELPNKVADLEKEISNLEILLG 58
   |||||

QY 61 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 104
   |||||
Db 59 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 102
   |||||

RESULT 9
US-10-299-636-67
; Sequence 67, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
```

```
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR FILING DATE: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 67
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (129)
; OTHER INFORMATION: Xaa at position 129 is unknown
; NAME/KEY: UNSURE
; LOCATION: (131)
; OTHER INFORMATION: Xaa at position 131 is unknown
; US-10-299-636-67

Query Match          90.5%; Score 465; DB 15; Length 233;
Best Local Similarity 92.3%; Pred. No. 2.9e-31;
Matches 96; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 LAKQTELEKLLDSDPEGKTQDELKAEAEALDKKADELPNKVADLEKEISNLEILLG 60
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 53 LAQKPTGLEKLLDSDPEGKTQDELKAEAEALDKKADELPNKVADLEKEISNLEILLG 112

Qy 61 GADSEDDTAALPNKATKKALEKTKOKELDAALNELGPDGDEEE 104
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 113 GADSEDDTAALPNKATKKALEKTKOKELDAALNELGPDGDEEE 156

RESULT 10
US-10-674-755-19
; Sequence 19, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-10-674-755-19

Query Match          73.5%; Score 378; DB 15; Length 80;
Best Local Similarity 96.2%; Pred. No. 1.7e-24;
Matches 77; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 25 LDKAEAEALDKKADELPNKVADLEKEISNLEILLGGADSEDDTAALPNKATKKALEK 84
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 LDKAEAEALDKKADGUPNKVSDLEKEISNLEILLGGADSEDDTAALPNKATKKALEK 60

Qy 85 TQKELDAALNELGPDGDEEE 104
```

```
Db 61 TQKELDAALNELGPDGDEEE 80
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
US-10-674-755-23
; Sequence 23, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-10-674-755-23

Query Match          62.1%; Score 319; DB 15; Length 108;
Best Local Similarity 64.8%; Pred. No. 2.1e-19;
Matches 70; Conservative 13; Mismatches 21; Indels 4; Gaps 2;

Qy 1 LAKQTELEKLLDSDPEGKTQDELKAEAEALDKKADELPNKVADLEKEISNLEILLG 60
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 LKAEAELENLLSTLDPEGKTQDELKAEAEALDKKADELPNKVADLEKEISNLEILLG 60

Qy 61 GADS---ED-DTAALPNKATKKALEKTKOKELDAALNELGPDGDEEE 104
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 DAETNNVEDYIKGLEAEIATKQAELEKTPKELDAALNELGPDGDEEE 108

RESULT 12
US-10-299-636-82
; Sequence 82, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 82
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-10-299-636-82

Query Match          62.1%; Score 319; DB 15; Length 211;
Best Local Similarity 64.8%; Pred. No. 4.6e-19;
Matches 70; Conservative 13; Mismatches 21; Indels 4; Gaps 2;

Qy 1 LAKQTELEKLLDSDPEGKTQDELKAEAEALDKKADELPNKVADLEKEISNLEILLG 60
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 25 LKAEAELENLLSTLDPEGKTQDELKAEAEALDKKADELPNKVADLEKEISNLEILLG 84
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:53:39 ; Search time 10.4 Seconds
(without alignments)
962.168 Million cell updates/sec

Title: US-10-674-755-20

Perfect score: 514

Sequence: 1 LAKKQTELEKLDLSDLPQEK.....TKKELDALNELGPDGDEE 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: piri:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	498	96.9	744	2	P95013
2	128	24.9	619	2	A97887
3	128	24.9	619	2	A41971
4	109.5	21.3	388	2	A46173
5	108	21.0	1837	2	T41023
6	107	20.8	385	2	T20410
7	106.5	20.7	1509	1	A27224
8	104.5	20.3	388	2	S52536
9	104.5	20.3	924	2	S06117
10	104.5	20.3	2007	1	B43402
11	104	20.2	281	2	F75216
12	103	20.0	1964	2	A59282
13	102.5	19.9	405	2	A33939
14	102.5	19.9	2139	2	T18296
15	101.5	19.7	1937	2	I38055
16	101	19.6	1175	2	C35815
17	101	19.6	1175	2	D35815
18	101	19.6	1201	2	A35815
19	101	19.6	1201	2	B35815
20	101	19.6	2385	2	A32491
21	101	19.6	2411	2	B32491
22	100.5	19.6	629	2	T44607
23	100.5	19.6	1938	1	A40997
24	100	19.5	259	2	D60110
25	100	19.5	318	2	T49167
26	99.5	19.4	397	2	H86754
27	99.5	19.4	1992	1	S02771
28	99	19.3	516	2	B84709
29	98.5	19.2	527	2	S33068

30 98.5 19.2 1940 2 A59287
31 98.5 19.2 1976 2 A59252
32 98 19.1 248 2 T26412
33 98 19.1 1169 2 A64505
34 98 19.1 1939 2 T18372
35 97.5 19.0 284 2 A45488
36 97.5 19.0 387 2 S57834
37 97.5 19.0 1119 2 T14321
38 97.5 19.0 1475 2 T33318
39 97 18.9 518 2 G84488
40 97 18.9 587 2 JC1419
41 97 18.9 1063 2 T18255
42 97 18.9 1179 2 F71190
43 96.5 18.8 281 2 A34787
44 96.5 18.8 284 1 TMRBA
45 96.5 18.8 284 2 A39816

ALIGNMENTS

RESULT 1

F95013

pneumococcal surface protein A [imported] - Streptococcus pneumoniae (strain TIGR4)

C;Species: Streptococcus pneumoniae

C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004

C;Accession: F95013

R;Tetzelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heidt

on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, I.

nson, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A;Reference number: A95000; MUID:21357209; PMID:11463916

A;Accession: F95013

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-744 <KUR>

A;Cross-references: UNIPROT:Q97T39; GB:AE005672; PIDN:AAK74303.1; PID:g14971584; GSPDB:G

R;Experimental source: strain TIGR4

C;Genetics:

A;Gene: SP0117

Query Match 96.9%; Score 498; DB 2; Length 744;
Best Local Similarity 98.1%; Pred. No. 2.2e-26;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDLSDLPQEKTDQELDKAEAEAEALDKADELPKNVADLEKEISNLEILLG 60

DB 346 LAKKQTELEKLDLSDLPQEKTDQELDKAEAEAEALDKADELPKNVADLEKEISNLEILLG 405

QY 61 GADSEDDTAALPNKLTAKAELEKTKQELDAALNELGPDGDEE 104

DB 406 GADSEDDTAALQNKLTAKAELEKTKQELDAALNELGPDGDEE 449

RESULT 2

A97887

surface protein pspA precursor [imported] - Streptococcus pneumoniae (strain R6)

C;Species: Streptococcus pneumoniae

C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004

C;Accession: A97887

R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczyk, L.; Burgett, S.; DeHoff, B.S.; Mc

e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; Mc

Y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A;Reference number: A97887; MUID:21429245; PMID:11544234

A;Accession: A97887

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-619 <KUR>

A;Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:AE007317; PIDN:AAK98925.1; PID:gl
C;Genetics:
A;Gene: pspA

Query Match 24.9%; Score 128; DB 2; Length 619;
Best Local Similarity 33.6%; Pred. No. 0.14;
Matches 38; Conservative 19; Mismatches 34; Indels 22; Gaps 3;

Qy 4 KQTELEKLDSLDPEGKTQDELKDE-----AEEAELDKKADELPNKVAD 47
Db 208 KIALENQVHLEQELKEIDEESEDYAKGFRAPLOSKLDKAKKL-SKLEELSDKIDE 266

Qy 48 LEKEISNLEILLGASDDT-----AALPNKLATKAELEKTKQKELDAALNE 95
Db 267 LDABIAKLEQLKAAEENNVEDYFKEGLEKTIAAKKAEELEKTEADLKAVNE 319

RESULT 3
A41971
surface protein pspA precursor - Streptococcus pneumoniae
N;Alternate names: pneumococcal surface protein A
C;Species: Streptococcus pneumoniae
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
A;Accession: A41971; A60282; A31134
R;Yother, J.; Briles, D.E.
J. Bacteriol. 174, 601-609, 1992
A;Title: Structural properties and evolutionary relationships of PspA, a surface protein
A;Reference number: A41971; MUID:92105030; PMID:1729249
A;Accession: A41971
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-619 <YOT>
A;Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:M74122; NID:gl53840; PIDN:AAA2701
A;Note: sequence extracted from NCBI backbone (NCBIN:75635, NCBIP:75636)
R;Talkington, D.F.; Crimmins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.
Infect. Immun. 59, 1285-1289, 1991
A;Title: A 43-kilodalton pneumococcal surface protein, PspA: isolation, protective ability
A;Reference number: A60282; MUID:91169598; PMID:2004810
A;Accession: A60282
A;Molecule type: protein
A;Residues: 32-76 <TAL>
A;Experimental source: strain JY2008
C;Genetics:
A;Gene: pspA
F;1-31/Domain: signal sequence #status predicted <SIG>
F;32-619/Product: surface protein pspA #status predicted <MAT>
F;411-430/Domain: cpl repeat homology <CP01>
F;431-450/Domain: cpl repeat homology <CP02>
F;451-470/Domain: cpl repeat homology <CP03>
F;471-490/Domain: cpl repeat homology <CP04>
F;491-510/Domain: cpl repeat homology <CP05>
F;511-530/Domain: cpl repeat homology <CP06>
F;531-550/Domain: cpl repeat homology <CP07>
F;551-570/Domain: cpl repeat homology <CP08>
F;571-591/Domain: cpl repeat homology <CP09>
F;592-611/Domain: cpl repeat homology <CP10>

Query Match 24.9%; Score 128; DB 2; Length 619;
Best Local Similarity 33.6%; Pred. No. 0.14;
Matches 38; Conservative 19; Mismatches 34; Indels 22; Gaps 3;

Qy 4 KQTELEKLDSLDPEGKTQDELKDE-----AEEAELDKKADELPNKVAD 47
Db 208 KIALENQVHLEQELKEIDEESEDYAKGFRAPLOSKLDKAKKL-SKLEELSDKIDE 266

Qy 48 LEKEISNLEILLGASDDT-----AALPNKLATKAELEKTKQKELDAALNE 95
Db 267 LDABIAKLEQLKAAEENNVEDYFKEGLEKTIAAKKAEELEKTEADLKAVNE 319

RESULT 4
A46173
MrpA protein - Streptococcus sp. (group A)

C;Species: Streptococcus sp.
C;Date: 21-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 30-May-1997
C;Accession: A46173
R;O'Toole, P.; Stenberg, L.; Rissler, M.; Lindahl, G.
Proc. Natl. Acad. Sci. U.S.A. 89, 8661-8665, 1992
A;Title: Two major classes in the M protein family in group A streptococci.
A;Reference number: A46173; MUID:92409576; PMID:1528877
A;Contents: group A
A;Accession: A46173
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-388 <OIT>
A;Note: sequence extracted from NCBI backbone (NCBIN:114063, NCBIP:114064)
C;Superfamily: M5 protein

Query Match 21.3%; Score 109.5; DB 2; Length 388;
Best Local Similarity 33.0%; Pred. No. 1.5;
Matches 36; Conservative 25; Mismatches 41; Indels 7; Gaps 4;

Qy 1 LAKKQTE---LEKLLDSLDPE-GKTQDELQ-KEAEAEALDKKADELPNKVADLEKEISNL 55
Db 182 IAKLQSEATLENLGSAKRELTELQAKLDATATAEKAKUESQVTTLENLGSAKRELTDL 241

Qy 56 EILLGGADSEDDTAALPNKLATKAELEKTKQKELDAALNELGPDGDEEE 104
Db 242 QAKLDAANAEE--KLQSQATLEKQLKELATKELADLQAKLAATNQEKE 288

RESULT 5
T41023
probable nuclear pore complex-associated protein - fission yeast (Schizosaccharomyces po
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T41023
R;Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, June 1998
A;Reference number: Z21965
A;Accession: T41023
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-1837 <MUR>
A;Cross-references: UNIPROT:O74424; EMBL:AL023860; PIDN:CAA19588.1; GSPDB:GN000068; SPDB:5
A;Experimental source: strain 972h-; cosmid c162
C;Genetics:
A;Gene: SPDB:SPCC162.08C
A;Map position: 3

Query Match 21.0%; Score 108; DB 2; Length 1837;
Best Local Similarity 31.1%; Pred. No. 9;
Matches 37; Conservative 21; Mismatches 35; Indels 26; Gaps 5;

Qy 1 LAKKQTELEKL-LDSLDPEGKTQ-DELDKEAEAEALDKKADELPNKVADLEKEISNL--- 55
Db 1500 LADSKNEHLQSEAVDADGKTSINLEKEITHELRSDEG--LVQQVQNLUSAEALAREH 1557

Qy 56 EILLGGADSEDDTAALPNKLATKK-----AELEKTKQKELDAALNE 95
Db 1558 SPTQGSLENADELARLSQLESTKQYYEKEKETELLAARSELVAEREKTKKEELNQLNE 1616

RESULT 6
T20410
hypothetical protein E02A10.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T20410
R;Thomas, K.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19271
A;Accession: T20410
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-385 <WIL>

A;Accession: S52536
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-388 <KAT>
A;Cross-references: UNIPROT:Q53474; GB:S75411; NID:g914107; PIDN:AAB33261.1; PID:g914109
C;Superfamily: M5 protein

Query Match 20.3%; Score 104.5; DB 2; Length 388;
Best Local Similarity 33.0%; Pred. No. 3.2;
Matches 36; Conservative 23; Mismatches 43; Indels 7; Gaps 4;

Qy 1 LAKQTE---LEKLSDLDPE-GKTQDEL-DKEAEALDKKADELPNKVADLKEISNL 55
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
182 IAKQSEANTLENLGSAKRELTDAQUDTATAEKALKESQVTTLENLGSAKRELTDL 241
:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:

Qy 56 ILGLGADSEDDTA-ALPNKLYATKKAELEKTKQKELDAALNELGPDGDDEE 104
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
242 QAKLDANAE--KAKLSQAALAEKLEATKKELADLAQAKLAATNQEKE 288
:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:

RESULT 9
S06117
myosin heavy chain, nonmuscle (clone lambda-FWHC) - chicken (fragment)
C;Species: Gallus gallus (Chicken)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: S06117
R;Katsuragawa, Y.; Yanagisawa, M.; Inoue, A.; Masaki, T.
Eur. J. Biochem. 184, 611-616, 1989
A;Title: Two distinct nonmuscle myosin-heavy-chain mRNAs are differentially expressed in
S.

A;Reference number: S06116; MUID:90032648; PMID:2806244
A;Accession: S06117
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-924 <KAT>
A;Cross-references: UNIPROT:Q02015; GB:X17590
A;Note: This translation is not annotated in GenBank entry GGMHCFPC, release 114
C;Superfamily: myosin heavy chain; myosin motor domain homology
F;1-303/Domain: myosin motor domain homology (fragment) <MMOT>

Query Match 20.3%; Score 104.5; DB 2; Length 924;
Best Local Similarity 30.3%; Pred. No. 7.7;
Matches 33; Conservative 26; Mismatches 43; Indels 7; Gaps 3;

Qy 1 LAKQTELEKLSDLD----PEGKTQDELDKAEAEALDKKADELPNKVADLKEISNL 56
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
557 LAKLNKGEMMTDLERLKKEETKQELEK--AKRKLDGETTDLQDQIAELQAIIEBK 614
:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:

Qy 57 ILGLGADSEDDTA-ALPNKLYATKKAELEKTKQKELDAALNELGPDGDDEE 104
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
615 IQLAKKEBELQALARGDGEAVQKNALKVIRELQIAELQEDLESEK 663
:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:

RESULT 10
B43402
myosin heavy chain-B, neuronal - chicken
N;Contains: myosin ATPase (EC 3.6.4.1)
C;Species: Gallus gallus (Chicken)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: B43402; A43402
R;Takahashi, M.; Kawamoto, S.; Adelstein, R.S.
J. Biol. Chem. 267, 17864-17871, 1992
A;Title: Evidence for inserted sequences in the head region of nonmuscle myosin specific
Yosin.

A;Reference number: A43402; MUID:92388144; PMID:1355479
A;Accession: B43402
A;Molecule type: mRNA
A;Residues: 1-2007 <TAK>
A;Cross-references: UNIPROT:Q02015; GB:M93676; NID:g212448; PIDN:AAA48988.1; PID:g212452
A;Note: The sequence of residues 212-221 and 632-652 and the corresponding nucleotide se
A;Accession: A43402
A;Molecule type: mRNA
A;Residues: 1-211;222-631;653-2007 <TA2>

C:Superfamily: myosin heavy chain; myosin motor domain homology
 F;91-780/Domain: myosin motor domain homology <MMOT>

Query Match 19.9%; Score 102.5; DB 2; Length 2139;
 Best Local Similarity 32.4%; Pred. No. 25;
 Matches 35; Conservative 16; Mismatches 36; Indels 21; Gaps 4;

QY 1 LAKKQTELEKLLSLDPGKTDQELQKEAEAEALDKADLPNKNVADLEKEISNLE---I 57
 DB 1507 VAKLNTQITKL-----TRDQSAEELNELSRKADKKKIKISELEQVNELESRPV 1557

QY 58 LLGGADS-----EDDTAALPNKLATKKA-----ELEKTKDELDAALNEL 96
 DB 1558 GTGNADENEIKIRDAIADLNKALEMKGVQNNQATNKLKAKNDL 1605

RESULT 15
 I38055
 myosin heavy chain, perinatal skeletal muscle - human
 N:Contains: myosin ATPase (EC 3.6.4.1)
 C:Species: Homo sapiens (man)
 C:Date: 17-May-1996 #sequence revision 17-May-1996 #text change 09-Jul-2004
 C:Accession: I38055; JH0154; S12459; S09332; A30220; S49478
 R:Jullian, E.H.; Kelly, A.M.; Pompidou, A.J.; Hoffman, R.; Schiaffino, S.; Stedman, H.H.
 Eur. J. Biochem. 230, 1001-1006, 1995
 A:Title: Characterization of a human perinatal myosin heavy-chain transcript.
 A:Reference number: I38055; MUID:95324556; PMID:7601129
 A:Accession: I38055
 A:Status: preliminary; translated from GB/EMBL/DDBY
 A:Molecule type: mRNA
 A:Residues: 1-1937 <RES>
 A:Cross-references: UNIPROT:P13535; EMBL:X238133; NID:9558668; PIDN:CAA86293.1; PID:9558668
 R:Karsch-Mizrachi, I.; Feghali, R.; Shows, T.B.; Leinwand, L.A.
 Gene 89, 289-294, 1990
 A:Title: Generation of a full-length human perinatal myosin heavy-chain-encoding cDNA.
 A:Reference number: JH0154; MUID:90323631; PMID:2373371
 A:Accession: JH0154
 A:Molecule type: mRNA
 A:Residues: 1-14, 'A', 16-859 <KAR>
 A:Cross-references: GB:Y00821
 A:Experimental source: skeletal muscle
 R:Bober, E.
 Submitted to the EMBL Data Library, January 1989
 A:Reference number: S12458
 A:Accession: S12459
 A:Molecule type: mRNA
 A:Residues: 502-1071, 'N', 1073-1250, 'DGG', 1253-1376, 'NT', 1379-1913, 'D', 1915-1937 <BOB>
 A:Cross-references: EMBL:X51592; NID:929465; PIDN:CAA35941.1; PID:929466
 A:Experimental source: clone gCMC-F
 R:Bober, E.; Buchberger-Seidl, A.; Braun, T.; Singh, S.; Goedde, H.W.; Arnold, H.H.
 Eur. J. Biochem. 189, 55-65, 1990
 A:Title: Identification of three developmentally controlled isoforms of human myosin heavy chain
 A:Reference number: S09331; MUID:90235862; PMID:1691980
 A:Accession: S09332
 A:Molecule type: mRNA
 A:Residues: 502-547, 'X', 549-617, 'X', 619-687, 'X', 689-757, 'X', 759-827, 'X', 829-897, 'X', 899-1376, 'NT', 1379-1386, 'X', 1388-1456, 'X', 1458-1526, 'X', 1528-1596, 'X', 1598-1666, 'X', 1668-1713, 'NT', 1715-1797, 1799-1937 <BOB>
 A:Cross-references: EMBL:X51592
 R:Feghali, R.; Leinwand, L.A.
 J. Cell Biol. 108, 1791-1797, 1989
 A:Title: Molecular genetic characterization of a developmentally regulated human perinatal myosin heavy chain
 A:Reference number: A30220; MUID:89234168; PMID:2715179
 A:Accession: A30220
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 860-969, 'Q', 971-1246, 'H', 1248-1260, 'G', 1262-1296, 'Q', 1298-1503, 'AH', 1506-1844, 'NT', 1846-1937 <BOB>
 A:Cross-references: GB:Y00821; NID:934863; PIDN:CAA68757.1; PID:934864
 C:Genetics:
 A:Gene: GDB:MYH8
 A:Cross-references: GDB:125267; OMIM:160741
 A:Map position: 17pter-17p12
 C:Superfamily: myosin heavy chain; myosin motor domain homology
 C:Keywords: actin binding; ATP; coiled coil; hydrolyase; muscle contraction; nucleotide binding

F;91-769/Domain: myosin motor domain homology <MMOT>
 F;181-188/Region: nucleotide-binding motif A (P-loop)
 F;551-588/Region: actin binding #status predicted
 F;658-680/Region: actin binding #status predicted
 F;842-1282/Region: S2 #status predicted
 F;698,708/Active site: Cys #status predicted

Query Match 19.7%; Score 101.5; DB 2; Length 1937;
 Best Local Similarity 31.1%; Pred. No. 26;
 Matches 37; Conservative 22; Mismatches 31; Indels 29; Gaps 5;

QY 1 LAKKQTELEKLLD-----SLDPGKTDQELD-----KEAEAEALDKAD--ELPNK 44
 DB 1023 LTKAKTKLQEQVDDLEGSLEQEKLRMDLERAKRKLEGLDKLAQESTMDNENKQQLDEK 1082

QY 45 VADLEKEISNLEILLGGADSEDDTAALPNKLATKKALEKTKDELDAALNELPGDGDDEE 103
 DB 1083 LEKKEFEISNLI-----SKIEDEQAV-----EIQLQKKIKELQARIELGEIEAE 1128

Search completed: June 21, 2005, 10:12:05
 Job time : 11.4 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:54:24 ; Search time 64.4163 Seconds
(without alignments)
826.751 Million cell updates/sec

Title: US-10-674-755-20

Perfect score: 514

Sequence: 1 LAKKQTELEKLLDLDPEK.....TKELDAALNELGPDGDEE 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	498	96.9	228	2 Q9L5B8	Q9L5B8 streptococc
2	498	96.9	235	2 Q9L582	Q9L582 streptococc
3	498	96.9	249	2 Q9L5D4	Q9L5D4 streptococc
4	498	96.9	252	2 Q9L583	Q9L583 streptococc
5	498	96.9	360	2 Q8KOK3	Q8KOK3 streptococc
6	498	96.9	429	2 Q9LAX7	Q9LAX7 streptococc
7	498	96.9	526	2 Q9LAX9	Q9LAX9 streptococc
8	498	96.9	608	2 Q8VQ55	Q8VQ55 streptococc
9	498	96.9	744	2 Q97T39	Q97T39 streptococc
10	493	95.9	249	2 Q9L5B7	Q9L5B7 streptococc
11	484	94.2	211	2 Q9L579	Q9L579 streptococc
12	484	94.2	241	2 Q9L580	Q9L580 streptococc
13	481	93.6	249	2 Q9L585	Q9L585 streptococc
14	481	93.6	256	2 Q9L590	Q9L590 streptococc
15	480	93.4	502	2 Q9LAX8	Q9LAX8 streptococc
16	478	93.0	242	2 Q9L562	Q9L562 streptococc
17	474	92.2	209	2 Q9L593	Q9L593 streptococc
18	456	88.7	222	2 Q9L584	Q9L584 streptococc
19	324	63.0	246	2 Q9L5B4	Q9L5B4 streptococc
20	320	62.3	479	2 Q9LAX2	Q9LAX2 streptococc
21	320	62.3	481	2 Q9LAX5	Q9LAX5 streptococc
22	319	62.1	107	2 Q8KOK2	Q8KOK2 streptococc
23	313	60.9	653	2 Q34097	Q34097 streptococc
24	308	59.9	213	2 Q8GNS7	Q8GNS7 streptococc
25	295	57.4	480	2 Q9LAX3	Q9LAX3 streptococc
26	250.5	48.7	211	2 Q8GNT0	Q8GNT0 streptococc
27	250.5	48.7	257	2 Q9L594	Q9L594 streptococc
28	243.5	47.4	227	2 Q9KGS0	Q9KGS0 streptococc
29	243.5	47.4	256	2 Q9L595	Q9L595 streptococc
30	243.5	47.4	461	2 Q9LAX6	Q9LAX6 streptococc
31	139	27.0	417	2 Q9LAY3	Q9LAY3 streptococc

32	130.5	25.4	225	2 Q9L591	Q9L591 streptococc
33	130	25.3	415	2 Q9LAY1	Q9LAY1 streptococc
34	129.5	25.2	222	2 Q9L577	Q9L577 streptococc
35	129.5	25.2	262	2 Q9L576	Q9L576 streptococc
36	129.5	25.2	415	2 Q9LAY7	Q9LAY7 streptococc
37	128.5	25.0	246	2 Q9L578	Q9L578 streptococc
38	128.5	25.0	416	2 Q9LAY8	Q9LAY8 streptococc
39	128	24.9	619	2 Q54972	Q54972 streptococc
40	128	24.9	619	2 Q8DR10	Q8DR10 streptococc
41	127.5	24.8	237	2 Q9L592	Q9L592 streptococc
42	127.5	24.8	395	2 Q9LAY9	Q9LAY9 streptococc
43	127	24.7	249	2 Q9L575	Q9L575 streptococc
44	126.5	24.6	255	2 Q9L581	Q9L581 streptococc
45	126.5	24.6	255	2 Q9L5B6	Q9L5B6 streptococc

ALIGNMENTS

RESULT 1

Q9L5B8 PRELIMINARY; PRT; 228 AA.
ID Q9L5B8
AC Q9L5B8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PepA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=60;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Packham R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=60;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF253404; AAF67352.1; -
DR InterPro; IPR009053; Prefoldin.
FT NON_TER 1 1
FT NON_TER 228 228
SQ SEQUENCE 228 AA; 24430 MW; E6EAA953EC54EA0F CRC64;

Query Match 96.9%; Score 498; DB 2; Length 228;
Best Local Similarity 98.1%; Pred. No. 6.9e-25;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDLDPEKQDELKKEAEAEELDKADELPKNKVADELEKISNLEILLG 60
|||||

Db 46 LAKKQTELEKLLDLDPEKQDELKKEAEAEELDKADELPKNKVADELEKISNLEILLG 105
|||||

QY 61 GADSEDDTAALPNKATKAELEKTKELDAALNELGPDGDEE 104
|||||

Db 106 GADSEDDTAALQNKATKAELEKTKELDAALNELGPDGDEE 149
|||||

RESULT 2

Q9L582 PRELIMINARY; PRT; 235 AA.
ID Q9L582
AC Q9L582;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PepA (Fragment).

```

GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
RX NCBT_TaxID=1313;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=3;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669 (2000) .
RN [2]
RC SEQUENCE FROM N.A.
RP STRAIN=3;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255545; AAF68098.1; -.
FT NON_TER 1
FT NON_TER 235
SQ SEQUENCE 235 AA; 25424 MW; BFFB48C52CA8380 CRC64;

Query Match 96.9%; Score 498; DB 2; Length 235;
Best Local Similarity 98.1%; Pred. No. 7.1e-25;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LAKKQTELEKLLSLDPEGKTQDELKAEAEALDKADELPNKVADLEKEISNLEILLG 60
Db 56 LAKKQTELEKLLSLDPEGKTQDELKAEAEALDKADELPNKVADLEKEISNLEILLG 115

Qy 61 GADSEDDTAALPNKLTAKAELEKTOKELDAALNELGPDGDEEE 104
Db 116 GADSEDDTAALQNKLTAKAELEKTOKELDAALNELGPDGDEEE 159

RESULT 3
Q9L5D4 PRELIMINARY; PRT; 249 AA.
ID Q9L5D4
AC Q9L5D4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBT_TaxID=1313;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=SP195;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669 (2000) .
RN [2]
RC SEQUENCE FROM N.A.
RP STRAIN=SP195;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF252286; AAF69499.1; -.
FT HSSP; P04268; 1IC2.
FT NON_TER 1
FT NON_TER 249
SQ SEQUENCE 249 AA; 26821 MW; F8EA39225CF8D43F CRC64;

Query Match 96.9%; Score 498; DB 2; Length 249;
Best Local Similarity 98.1%; Pred. No. 7.5e-25;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 LAKKQTELEKLLSLDPEGKTQDELKAEAEALDKADELPNKVADLEKEISNLEILLG 60
Db 67 LAKKQTELEKLLSLDPEGKTQDELKAEAEALDKADELPNKVADLEKEISNLEILLG 126

Qy 61 GADSEDDTAALPNKLTAKAELEKTOKELDAALNELGPDGDEEE 104
Db 127 GADSEDDTAALQNKLTAKAELEKTOKELDAALNELGPDGDEEE 170

RESULT 4
Q9L583 PRELIMINARY; PRT; 252 AA.
ID Q9L583
AC Q9L583;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBT_TaxID=1313;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=127;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669 (2000) .
RN [2]
RC SEQUENCE FROM N.A.
RP STRAIN=127;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255544; AAF68097.1; -.
DR HSSP; P04268; 1IC2.
FT NON_TER 1
FT NON_TER 252
SQ SEQUENCE 252 AA; 27260 MW; 82DE13741F369CA2 CRC64;

Query Match 96.9%; Score 498; DB 2; Length 252;
Best Local Similarity 98.1%; Pred. No. 7.6e-25;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LAKKQTELEKLLSLDPEGKTQDELKAEAEALDKADELPNKVADLEKEISNLEILLG 60
Db 73 LAKKQTELEKLLSLDPEGKTQDELKAEAEALDKADELPNKVADLEKEISNLEILLG 132

Qy 61 GADSEDDTAALPNKLTAKAELEKTOKELDAALNELGPDGDEEE 104
Db 133 GADSEDDTAALQNKLTAKAELEKTOKELDAALNELGPDGDEEE 176

RESULT 5
Q8KQK3 PRELIMINARY; PRT; 360 AA.
ID Q8KQK3
AC Q8KQK3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Pneumococcal surface protein A (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBT_TaxID=1313;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=259/98;
RX MEDLINE=22170754; PubMed=12183557;

```


DT	01-OCT-2000	(TREMBlrel. 15, Created)
DT	01-OCT-2000	(TREMBlrel. 15, Last sequence update)
DT	01-MAR-2004	(TREMBlrel. 26, Last annotation update)
DE	PcpA	(Fragment).
GN	Name=pspA;	
OS	Streptococcus pneumoniae.	
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;	
OC	Streptococcus.	
OX	NCBI_TaxID=1313;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=EF3296;	
RX	MEDLINE=20448953; PubMed=10992499;	
RA	DOI=10.1128/JIAI.68.10.5889-5900.2000;	
RR	Hollingshead S.K., Becker R., Briles D.E.;	
RT	"Diversity of PcpA: mosaic genes and evidence for past recombination	
RT	in Streptococcus pneumoniae";	
RL	Infect. Immun. 68:5889-5900(2000).	
DR	EMBL; AF071816; AAP27712.1; -.	
DR	HSSP; P04268; IIC2.	
DR	InterPro; IPRO11047; Quin_alc_DH_like.	
FT	NON TER 526 526	
SQ	SEQUENCE 526 AA; 58106 MW; 5F1F564A2CB678AE CRC64;	
 Query Match 96.9%; Score 498; DB 2; Length 526; Best Local Similarity 98.1%; Pred. No. 1.5e-24; Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
Qy	1	LAKQTETEKLKLDSPGEGTKQDELDKAEAEALDKADELPNKVADLEKEISNLLEILLG 60
Dd	346	LAKQTETEKLKLDSPGEGTKQDELDKAEAEALDKADELPNKVADLEKEISNLLEILLG 405
Qy	61	GADSEDDTAALPNKATKKAELKTKQELDAALNELGPDGDEE 104
Dd	406	GADSEDDTAALQNKLATKKAELKTKQELDAALNELGPDGDEE 449
 RESULT 8 Q8VQ55 PRELIMINARY; PRT; 608 AA. AC Q8VQ55; DT 01-MAR-2002 (TREMBlrel. 20, Created) DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update) DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update) DE Pneumococcal surface protein A (Fragment). GN Name=pspA; OS Streptococcus pneumoniae. OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; OC Streptococcus. OX NCBI_TaxID=1313; RN [1] RP SEQUENCE FROM N.A. RC STRAIN=KNIH1156; RA Lee K.J., Bae S.M., Chung K.S.; RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases. RL EMBL; AF460993; AAL67804.1; -. DR HSSP; P06653; IHGX. DR Pfam; PF01473; CW binding 1; 10. DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_2. FT NON_TER 1 1 FT NON TER 608 608 SQ SEQUENCE 608 AA; 67918 MW; 15F71BD62E297526 CRC64;		
 Query Match 96.9%; Score 498; DB 2; Length 608; Best Local Similarity 98.1%; Pred. No. 1.8e-24; Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
Qy	1	LAKQTETEKLKLDSPGEGTKQDELDKAEAEALDKADELPNKVADLEKEISNLLEILLG 60
Dd	222	LAKQTETEKLKLDSPGEGTKQDELDKAEAEALDKADELPNKVADLEKEISNLLEILLG 281
Qy	61	GADSEDDTAALPNKATKKAELKTKQELDAALNELGPDGDEE 104

```
Db 282 GADSEDDTAALQNKLATKKAEELEKTQKELDAALNELPGDGDDEE 325
RESULT 9
Q97T39 PRELIMINARY; PRT; 744 AA.
AC Q97T39;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pneumococcal surface protein A.
GN OrderedLocNames=Sp0117;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916; DOI=10.1126/science.1061217;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S.N., Heidelberg J.F., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E.K., Khouri H.M., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S.V., Dickinson T.,
RA Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae."
RL Science 293:498-506 (2001).
DR EMBL; AE007328; AAK74303.1; -.
DR PIR; F95013; F95013.
DR HSP; P06653; 1HCX.
DR TIGR; SP0117; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam; PF01473; CW binding 1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_2.
KW Complete proteome.
SQ SEQUENCE 744 AA; 82764 MW; 20EASE8E7911EPD5 CRC64;

Query Match 96.9%; Score 498; DB 2; Length 744;
Best Local Similarity 98.1%; Pred. No. 2.2e-24;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LAKQTELEKLLDSLDPEGKTQDELKAEAEALDKADELPNKVADLEKEISNLEILLG 60
Db 346 LAKQTELEKLLDSLDPEGKTQDELKAEAEALDKADELPNKVADLEKEISNLEILLG 405

Qy 61 GADSEDDTAALPNKLTAKKAELEKTQKELDAALNELPGDGDDEE 104
Db 406 GADSEDDTAALQNKLATKKAEELEKTQKELDAALNELPGDGDDEE 449

RESULT 10
Q9L5B7 PRELIMINARY; PRT; 249 AA.
AC Q9L5B7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=50;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones."
RL J. Clin. Microbiol. 38:3663-3669 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=20;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255548; AAF68101.1; -.
DR HSP; P04268; 1IC2.
FT NON_TER 1
FT NON_TER 231
SQ SEQUENCE 231 AA; 24990 MW; A7731F3A46460186 CRC64;

Query Match 94.2%; Score 484; DB 2; Length 231;
Best Local Similarity 95.3%; Pred. No. 5.6e-24;
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LAKQTELEKLLDSLDPEGKTQDELKAEAEALDKADELPNKVADLEKEISNLEILLG 60
Db 76 LAKQTELEKLLDNLDPEGKTQDELKAEAEALDKADELPNKVADLEKEISNLEILLG 135

Qy 61 GADSEDDTAALPNKLTAKKAELEKTQKELDAALNELPGDGDDEE 104
Db 136 GADPEDDTAALQNKLATKKAEELEKTQKELDAALNELPGDGDDEE 179

RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones."
RL J. Clin. Microbiol. 38:3663-3669 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=20;
RA Beall B.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255548; AAF68101.1; -.
DR HSP; P04268; 1IC2.
FT NON_TER 1
FT NON_TER 231
SQ SEQUENCE 231 AA; 24990 MW; A7731F3A46460186 CRC64;

Query Match 94.2%; Score 484; DB 2; Length 231;
Best Local Similarity 95.3%; Pred. No. 5.6e-24;
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LAKQTELEKLLDSLDPEGKTQDELKAEAEALDKADELPNKVADLEKEISNLEILLG 60
Db 76 LAKQTELEKLLDNLDPEGKTQDELKAEAEALDKADELPNKVADLEKEISNLEILLG 135

Qy 61 GADSEDDTAALPNKLTAKKAELEKTQKELDAALNELPGDGDDEE 104
Db 136 GADPEDDTAALQNKLATKKAEELEKTQKELDAALNELPGDGDDEE 179
```

```
RESULT 12
Q9L580 PRELIMINARY; PRT; 241 AA.
ID Q9L580;
AC Q9L580;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=121;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=121;
RA Beall B.W.;
RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255547; AAF68100.1; -.
DR HSSP; P04268; IIC2.
FT NON_TER 1 1
FT NON_TER 241 241
SQ SEQUENCE 241 AA; 26038 MW; BB87E1A4C25FA669 CRC64;

Query Match 94.2%; Score 481; DB 2; Length 241;
Best Local Similarity 95.2%; Pred. No. 5.8e-24;
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDSDLDPGKTDQELDKAEAEALDKADELPNKVADLKEISNLEILG 60
Db 77 LAKKQTELEKLLDSDLDPGKTDQELDKAEAEALDKADELPNKVADLKEISNLEILG 136

QY 61 GADSEDDTAAALPNKLTAKAELEKTKQELDAALNELGPGDDEE 104
Db 137 GADPEDDTAALQNKLTATTKAELEKTKQELDAALNELGPGDDEE 180

RESULT 13
Q9L585 PRELIMINARY; PRT; 249 AA.
ID Q9L585;
AC Q9L585;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=18;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=18;
RA Beall B.W.;
RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

QY 1 LAKKQTELEKLLDSDLDPGKTDQELDKAEAEALDKADELPNKVADLKEISNLEILG 60
Db 77 LAKKQTELEKLLDSDLDPGKTDQELDKAEAEALDKADELPNKVADLKEISNLEILG 136

QY 61 GADSEDDTAAALPNKLTAKAELEKTKQELDAALNELGPGDDEE 104
Db 137 GADPEDDTAALQNKLTATTKAELEKTKQELDAALNELGPGDDEE 180

RESULT 14
Q9L590 PRELIMINARY; PRT; 256 AA.
ID Q9L590;
AC Q9L590;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP193;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SP193;
RA Beall B.W.;
RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF254259; AAF68094.1; -.
DR HSSP; P04268; IIC2.
FT NON_TER 1 1
FT NON_TER 256 256
SQ SEQUENCE 256 AA; 27738 MW; 7F05351559AD9238 CRC64;

Query Match 93.6%; Score 481; DB 2; Length 256;
Best Local Similarity 95.2%; Pred. No. 9.6e-24;
Matches 99; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDSDLDPGKTDQELDKAEAEALDKADELPNKVADLKEISNLEILG 60
Db 66 LAKKQTELEKLLDSDLDPGKTDQELDKAEAEALDKADELPNKVADLKEISNLEILG 125

QY 61 GADSEDDTAAALPNKLTAKAELEKTKQELDAALNELGPGDDEE 104
Db 126 GADPEDDTAALQNKLTATTKAELEKTKQELDAALNELGPGDDEE 169

RESULT 15
Q9LAX8 PRELIMINARY; PRT; 502 AA.
ID Q9LAX8;
AC Q9LAX8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
```

```
GN Name=pspa;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG8090;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071817; AAF27713.1; -.
DR HSSP; O15813; 1D7M.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 502
SQ SEQUENCE 502 AA; 55018 MW; 4E73D477CAE79B40 CRC64;

Query Match      93.4%; Score 480; DB 2; Length 502;
Best Local Similarity 94.2%; Pred. No. 2.le-23;
Matches 98; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDSDPEGKTQDELDKAEAEALDKKADLPNKVADLEKEISNLEILLG 60
Db 347 LAKKQTELEKLLDNLDPGKTQDELDKAEAEALDKKADLPNKVADLEKEISNLEILLG 406

QY 61 GADSEDDTAALPNKATKKALEKTKQKELDAALNELGPDGDEE 104
Db 407 GADPEDDTAALQNKATKKALEKTKQKELDAALNELGPDGDEE 450
```

Search completed: June 21, 2005, 10:22:14
Job time : 64.4163 secs